



Four Ways Predictive Analytics Leverages Social Media

CASACT Ratemaking &
Product Management
Seminar
March 2014

Eric Siegel, Ph.D.
Founder, Predictive Analytics World
Author, *Predictive Analytics*

Eric Siegel, Ph.D.



PREDICTIVE ANALYTICS Times
The predictive analytics professionals' premier resource



BOOK:

Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die
www.thepredictionbook.com

CONFERENCE:

Predictive Analytics World
San Francisco, Chicago, Boston, Washington DC, Toronto, Berlin, and London
www.predictiveanalyticsworld.com

ONLINE PORTAL AND NEWS SITE:

Predictive Analytics Times
www.predictiveanalyticstimes.com

CONFERENCE:

Text Analytics World
San Francisco and Boston
www.textanalyticsworld.com

Online training:

"Predictive Analytics Applied" - *View it on-demand*
www.businessprediction.com



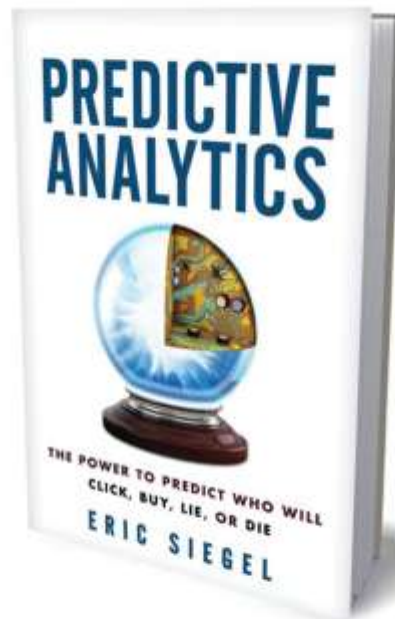
CHICAGO

June 17-18, 2014



BOSTON

October 6-7, 2014



www.ThePredictionBook.com

Agenda

- 1) Driving treatment selection with PA
- 2) Four ways PA leverages social media



Predictive Analytics:



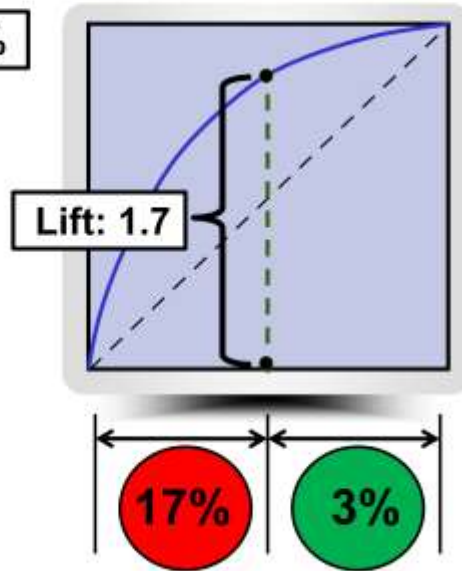
(data)

*Technology that learns from **experience** to predict the outcome/behavior of individuals ... in order to drive better decisions.*



Risk Detection

High claims: 10%



“ *It's not exclusively about predictive analytics and technical pricing models. We have to thoughtfully consider the blend between art and science and the effective/pragmatic implementation of the tools that we build. The “art” refers to the expertise and experience of the underwriter, and the “science” involves model development and the technology required to deploy them. Taken together, the models act as a tool to inform and support our underwriters in risk selection and pricing decisions.* ”

David Zona
Chief Underwriting Officer, Fireman's Fund

Predictive Analytics World



Predictive Analytics for Treatment Selection

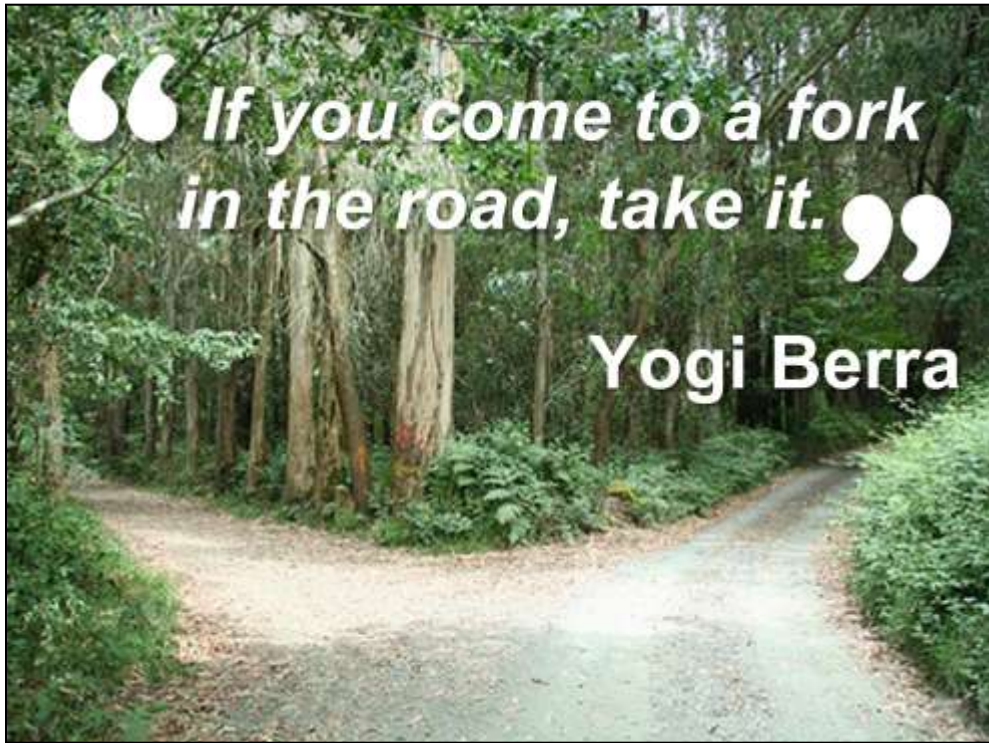
- 1) Pricing
- 2) Selection
- 3) Marketing
- 4) Fraud detection
- 5) Fast-tracking claims

Intervention

Predictive Analytics World







*“If you come to a fork
in the road, take it.”*

Yogi Berra





The Art Institutes logo: **AI The Art Institutes™**
America's Leader in Creative Education

The Art Institutes is an established system of schools located in major cities throughout North America. The Art Institutes offer master's, bachelor's and associate's degrees, diplomas and certificate programs that specialize in the creative and applied arts, including design, media arts, fashion and culinary arts.

Our curriculum at The Art Institutes is carefully scrutinized and developed by industry leaders and prominent employers who tell us what skills are needed in the workplace. This means that your education is relevant, practical, and hands-on.


Please enter your phone number * () -

I am interested in starting *

High school graduation date *

Note: To ensure a personalized response, we will forward your contact info, academic qualifications, and preferences (as noted in your [Client] profile) to The Art Institutes.

Yes, please have someone from The Art Institutes Admissions Office contact me!

Predictive Analytics World 

An example interstitial promotion. If the user accepts the offer, he/she is allowing the host to pass profile information directly to the sponsor (in addition to the fields shown).

A-B test deployment

- A. Legacy system based on acceptance rates across users
- B. Model-based ad selection

Results

25% increased "take" rate
3.6–5% revenue increase

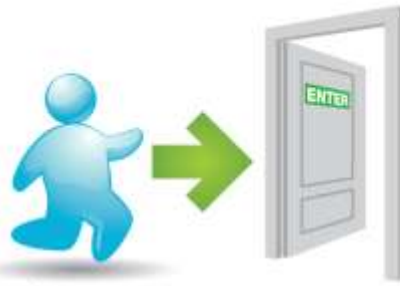
\$1 million every 19 months in additional revenue,
given the existing \$1.5 million monthly revenue

A few additional percentage points can be tough to get, in the face of fairly adept existing systems, but can make a big difference. Consider the insurance business, where predictive analytics aims to reduce the loss ratio by 2 to 5 points beyond that attained by standard actuarial methodology, or the engineering of jet engines, where a 1% increase in efficiency would be a huge bite out of annual fuel consumption.

The revenue results above are for *interstitial* ads only; many more ads are embedded within functional product web pages, and could also be targeted with only a slight alteration to the analytical system and deployment integration developed for this project.

The large 25% increase in acceptance rates means formerly less "popular" ads are now being given a better chance, leading to success; these sponsors are likely to appreciate the increase in customer leads now coming from advertising with the client.

Likewise, user satisfaction is likely higher, since users are seeing more ads in which they are provably more interested.



**1) Response modeling
for acquisition**



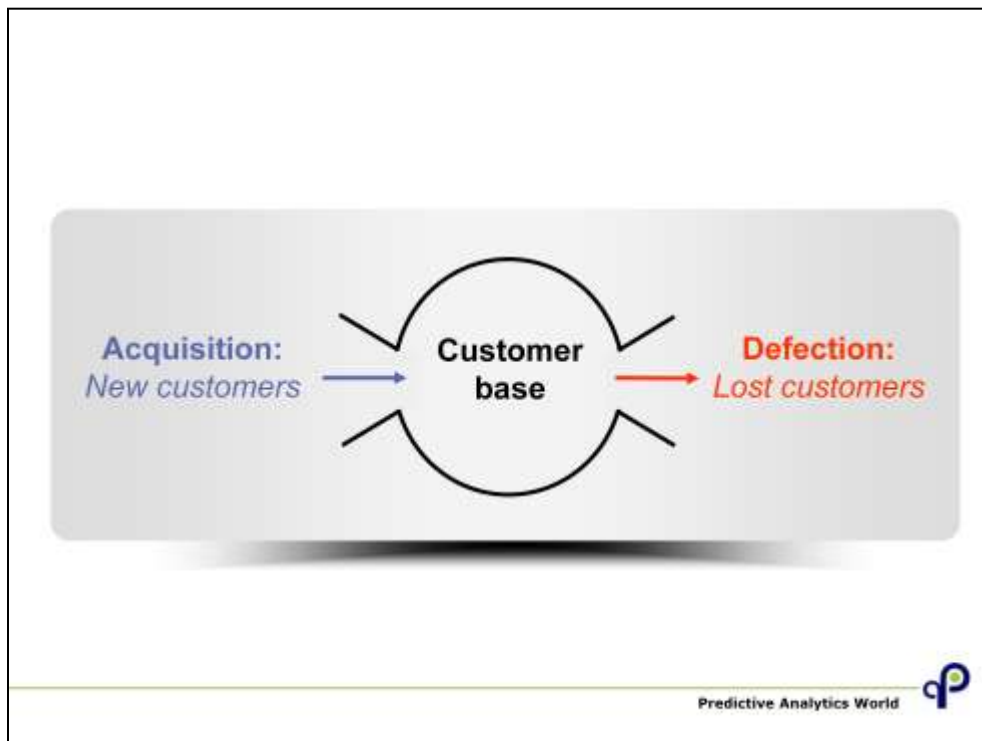
**2) Churn modeling
for retention**



3) Which recommendations, ad, content



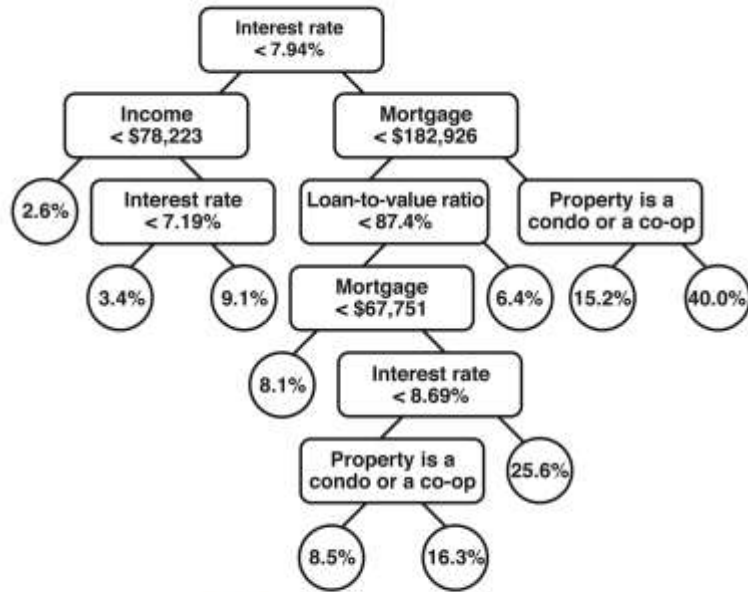
**2) Churn modeling
for retention**



The only way to target a retention campaign precisely where it's needed is with predictive scores that earmark which customers are most likely to leave.

Targeted retention is often the lowest hanging fruit among prospective applications of predictive analytics at an organization.

For more information, see Chapter 7 of the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" (<http://www.thepredictionbook.com>).



**RENEW YOUR
CELL PHONE
CONTRACT...**



**AND
GET A FREE
PHONE!**



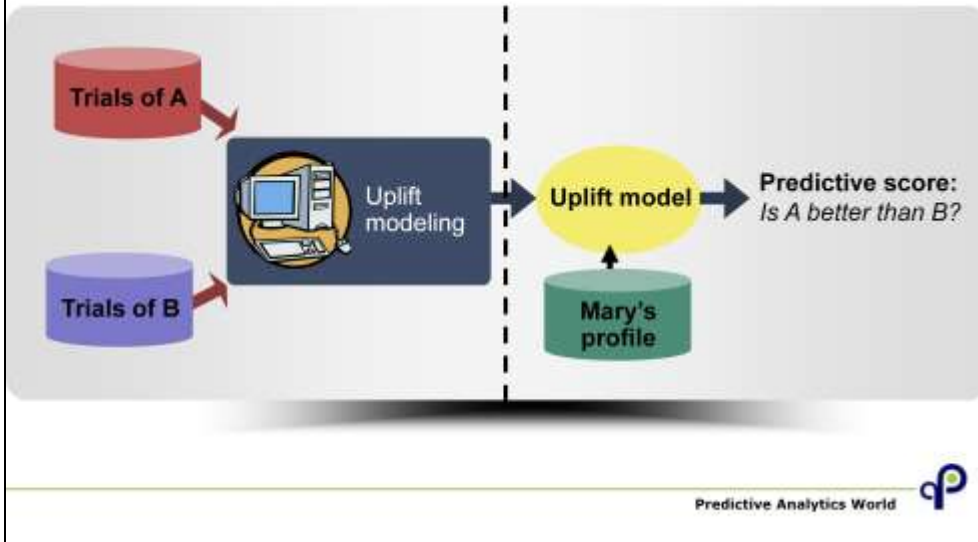


Treatment/Control
+
Modeling

PLACEBO

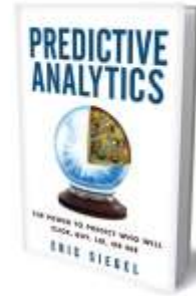
Predictive Analytics World

Uplift Modeling



Chapter 7: Persuasion by the Numbers

How Telenor, U.S. Bank, and the Obama Campaign Engineered Influence



White paper:

Uplift Modeling: Predictive Analytics Can't Optimize Marketing Decisions Without It

www.pawcon.com/signup-uplift-whitepaper.php



(This paper in turn references all the core technical papers on this topic.)

Free white paper: www.predictiveanalyticsworld.com/signup-uplift-whitepaper.php



Four Ways Predictive Analytics Leverages Social Media



SOCIAL + PREDICTIVE = ?



- *Who you know*
- *What you "Like"*

Predictive
Model



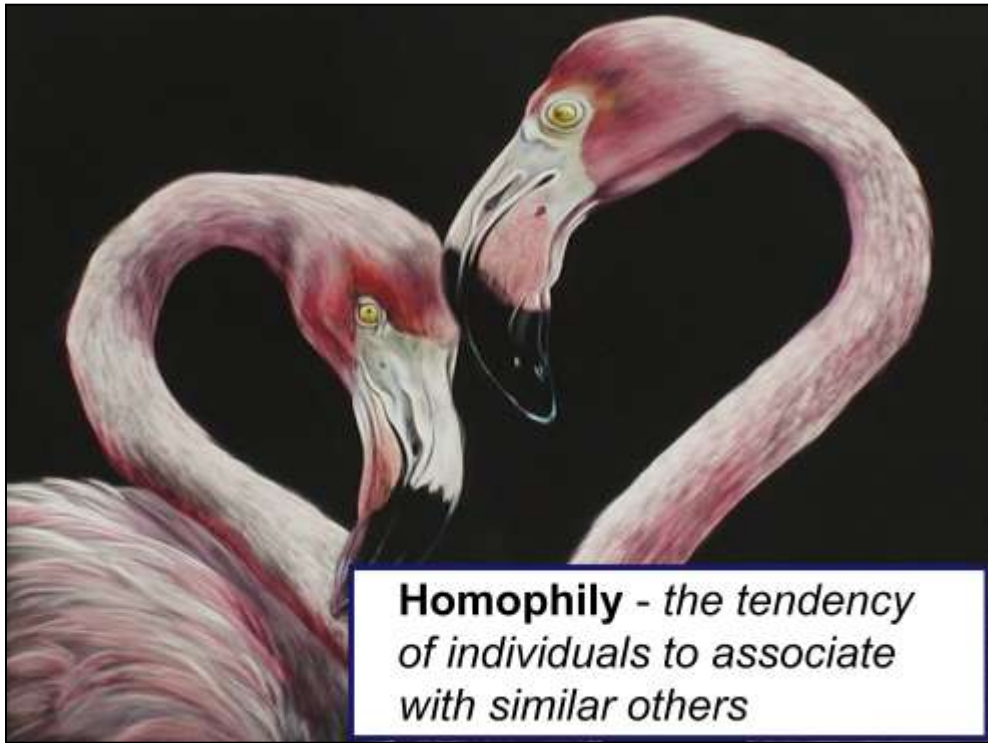
Predictive
Model

- *Who you know*
- *What you want*

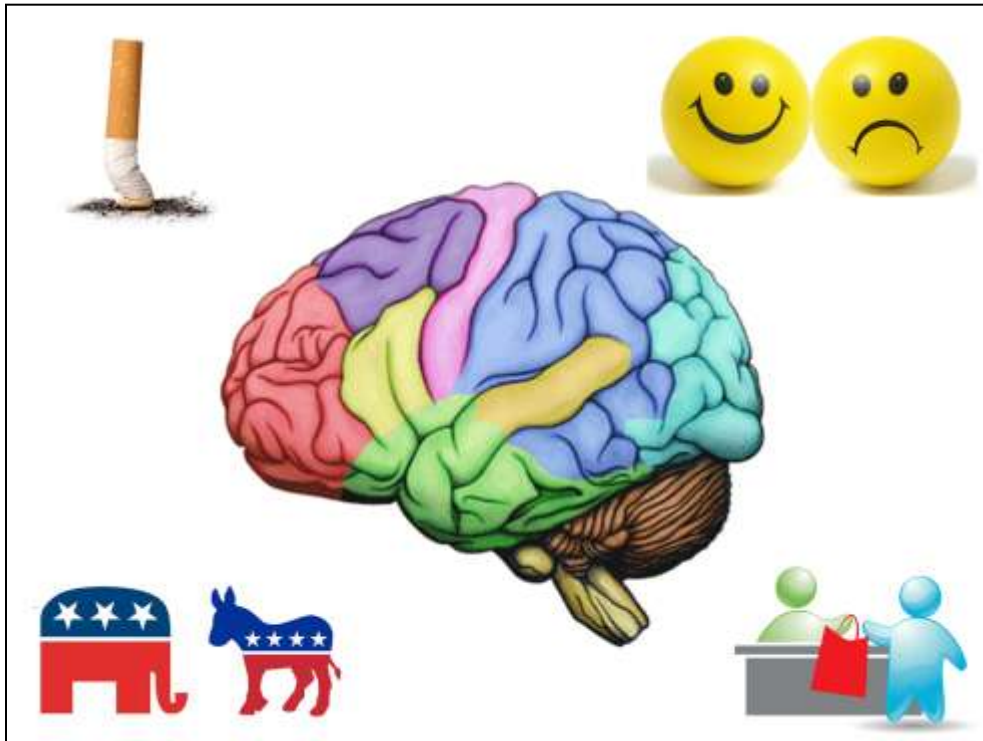
Four Ways Predictive Analytics Leverages Social Media

- 1. Predict behavior from who you know**
- 2. Predict behavior from social activity**
- 3. Predict what you like**
- 4. Predict who you know and your influence**





Homophily - *the tendency of individuals to associate with similar others*



Positive health habits are contagious. If you quit smoking, your close contacts become 36 percent less likely to smoke. Your chance of becoming obese increases by 57 percent if you have a friend who becomes obese.

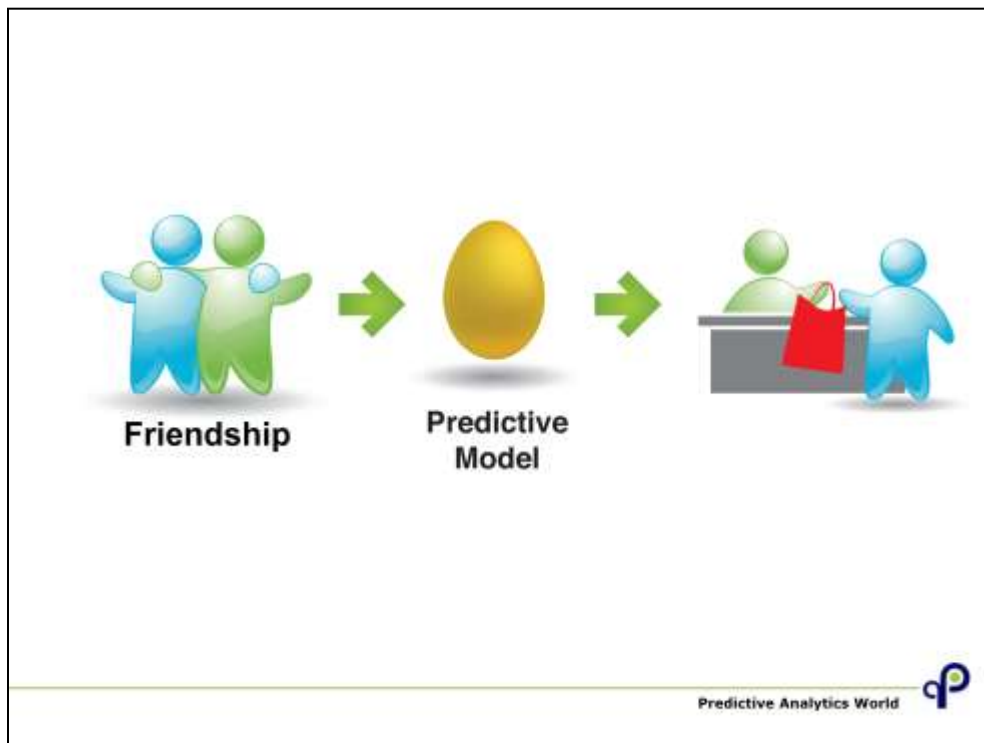
Happiness is contagious. Each additional Facebook friend who is happy increases your chances of being happy by roughly 9 percent. (Harvard University; "Waves of happiness . . . spread throughout the network.")

The Social Effect



The Social Effect





Also applies to fraud detection. If your contacts commit fraud, so might you.

In fact, one fraud scheme can't be detected without social data. A group of criminals open financial accounts that improve their respective credit ratings by transferring funds among themselves. Since the money transfers take place only between these accounts, the fraudsters need not spend any real money in conducting these transactions; they play their own little zero-sum game. Once each account has built up its own supposedly legitimate record, they strike, taking out loans, grabbing the money, and running. These schemes can be detected only by way of social data to reveal that the network of transactors is a closed group.

Major N. American Telecom
700% more likely to cancel if someone in your network does

Optus (Australian telecom)
Doubled churn model performance with *social data*



Predictive Analytics World 

If your friend's defect, you're much more likely to as well.

Friends stick to the same cell phone company. If you switch wireless carriers, your contacts are in turn up to seven times more likely to follow suit.

“Birds of a feather use the same phone service provider”

<http://blog.summation.net/2009/11/birds-of-a-feather-use-the-same-phone-service-provider.html>

“The Social Effect: Predicting Telecom Customer Churn with Call Data”

<http://www.predictiveanalyticsworld.com/sanfrancisco/2010/agenda.php#day1-12>

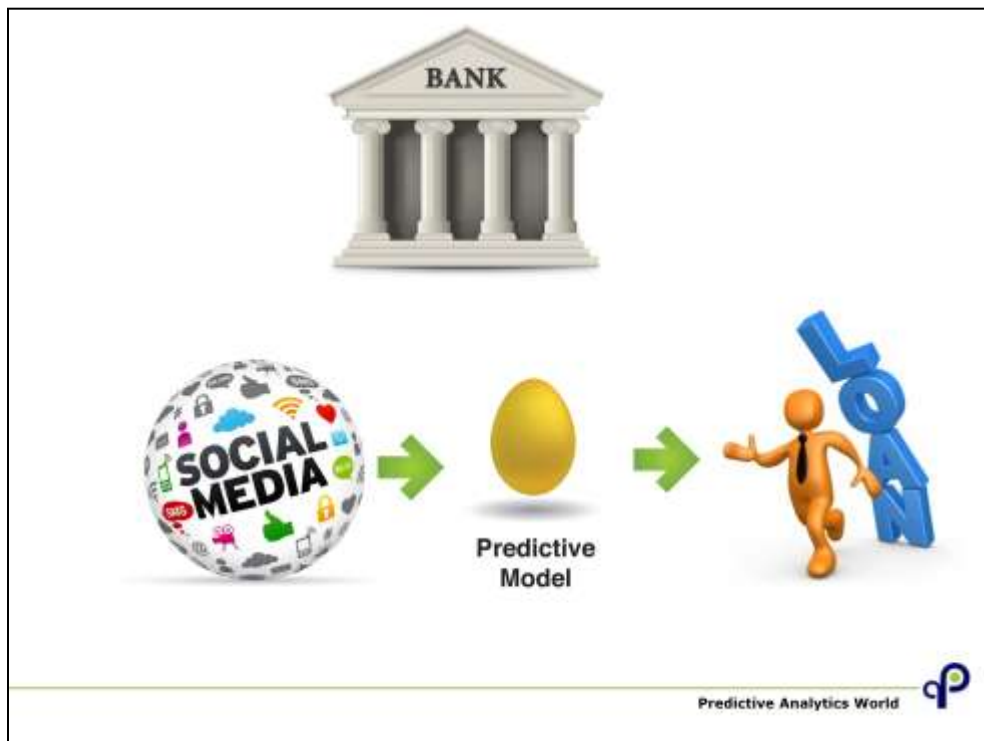
Four Ways Predictive Analytics Leverages Social Media

1. Predict behavior from who you know
2. **Predict behavior from social activity**
3. Predict what you like
4. Predict who you know and your influence

Predictive Analytics World



Related to this topic, cf. the prediction of protected classes from Facebook likes:
<http://www.cam.ac.uk/research/news/digital-records-could-expose-intimate-details-and-personality-traits-of-millions>



<http://mashable.com/2011/10/07/social-media-privacy-banks/>

Like IRS auditing and bank credit scoring, marketing is making moves to mine your social media activity



U.S. News & World Report highlighted some more encroaching news this past week in the article "IRS High-Tech Tools Track Your Digital Footprints." The IRS is "collecting a lot more than taxes this year," the report explained. "It's also acquiring a huge volume of personal information on taxpayers" digital activities, from eBay auctions to Facebook posts and, for the first time ever, credit card and e-payment transaction records." It added, "The agency reveals little about how it will employ its vast, new network scanning powers - sweeping changes being implemented with little public discussion or clear guidelines."

<http://money.usnews.com/money/personal-finance/mutual-funds/articles/2013/04/04/irs-high-tech-tools-track-your-digital-footprints>

<http://www.predictiveanalyticsworld.com/gov/2012/agenda.php#day1-115>



<http://www.ammoland.com/2013/04/irs-now-robo-audits-your-spending/>

Chuck Norris

ACTION JEANS

**"Won't bind
your legs"**



**PRICE
REDUCTION**
\$19⁹⁵
Adult Sizes

FREE CATALOG

of Century Martial Art products,
including: Chinese Kung Fu uniforms
and costumes, karate gi, taekwon
do uniforms, MMA, jiu-jitsu,
Brazilian Jiu-jitsu, BJJ, and
other martial arts gear.

Developed by Chuck Norris for
stunt fighting in action movies. These great
looking western style jeans have a unique
hidden zipper which allows greater
movement without binding or ripping.

Satisfaction Guaranteed. Check the fit and feel of
these great looking jeans. If you are not pleased, return them
within 14 days, postage prepaid, in new condition, for full
refund of purchase price or size adjustment.



**Best-Price Lifetime Guar-
antee.** If they ever wear
out, rip, fray, or tear, return
them (postage prepaid)
for a new pair at half the
retail price at the time of
return.



**Century Martial Art
Supply, Inc.**

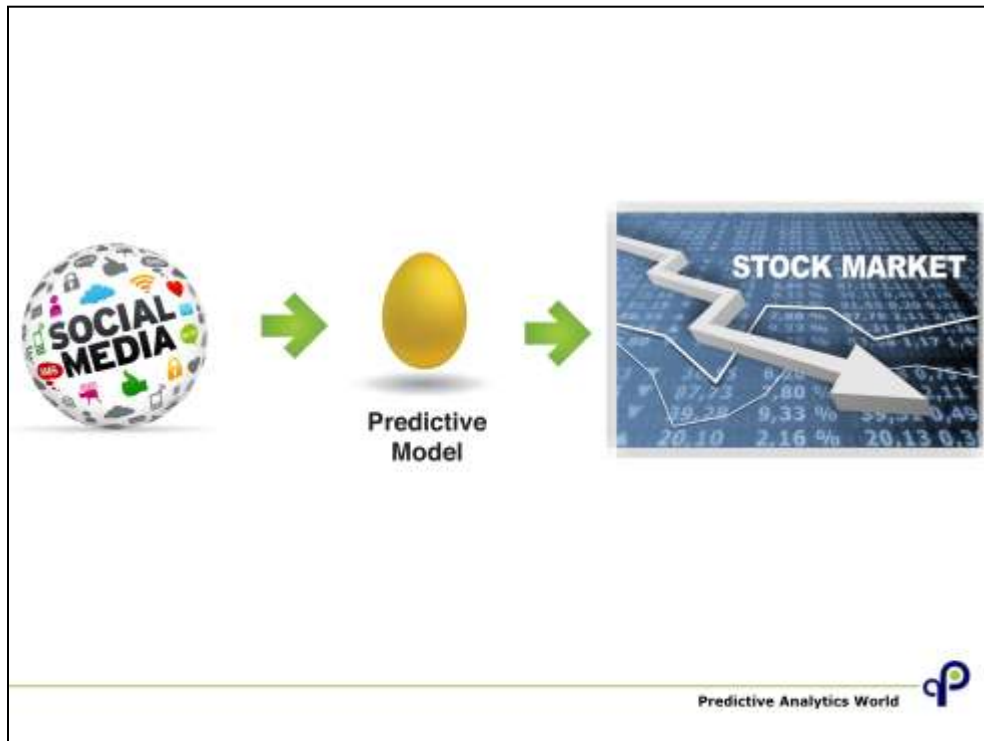
1102 National Blvd., Oklahoma City, OK 73110 1-800-625-2787

DEALER INQUIRES INVITED

ACTION JEANS ORDER FORM

Name _____
Address _____
City _____





<http://money.usnews.com/money/personal-finance/mutual-funds/articles/2013/04/04/irs-high-tech-tools-track-your-digital-footprints>

<http://www.predictiveanalyticsworld.com/gov/2012/agenda.php#day1-115>

<http://mashable.com/2011/10/07/social-media-privacy-banks/>



Social media (blogs, etc.) predicts the stock market, Product sales, Top music hits, Movie box-office revenue, Academy Award and Grammy winners, Elections, and Unemployment statistics

<http://money.usnews.com/money/personal-finance/mutual-funds/articles/2013/04/04/irs-high-tech-tools-track-your-digital-footprints>

<http://www.predictiveanalyticsworld.com/gov/2012/agenda.php#day1-115>

<http://mashable.com/2011/10/07/social-media-privacy-banks/>



What you like



Predictive
Model



Predictive
Model

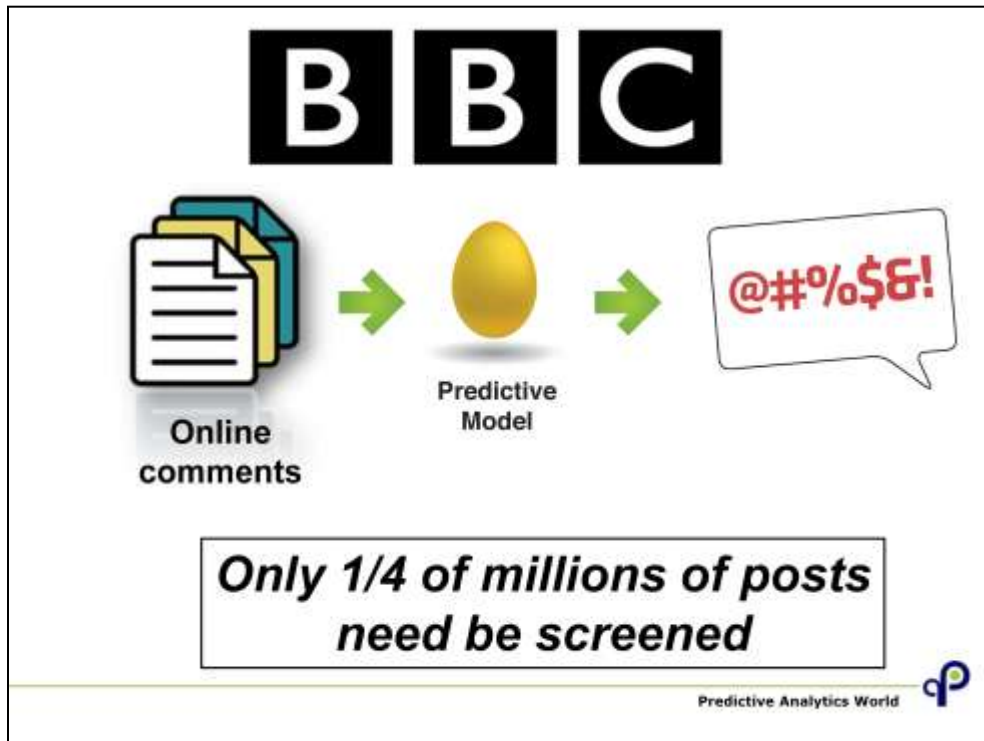


What you like/want

Four Ways Predictive Analytics Leverages Social Media

1. Predict behavior from who you know
2. Predict behavior from social activity
- 3. Predict what you like**
4. Predict who you know and your influence





The British Broadcasting Company predicts which comments will be approved for posting on its web pages so that only one-quarter of the millions of attempted posts need be screened by human moderators.

Text Analytics



Social Sentiment from Text

1. *Inappropriateness*
2. *Insults*
3. *Sarcasm*
4. *Politeness*
5. *Discontent*
6. *Anxiety*
7. *Love*
8. *Correctness*

Predictive Analytics World



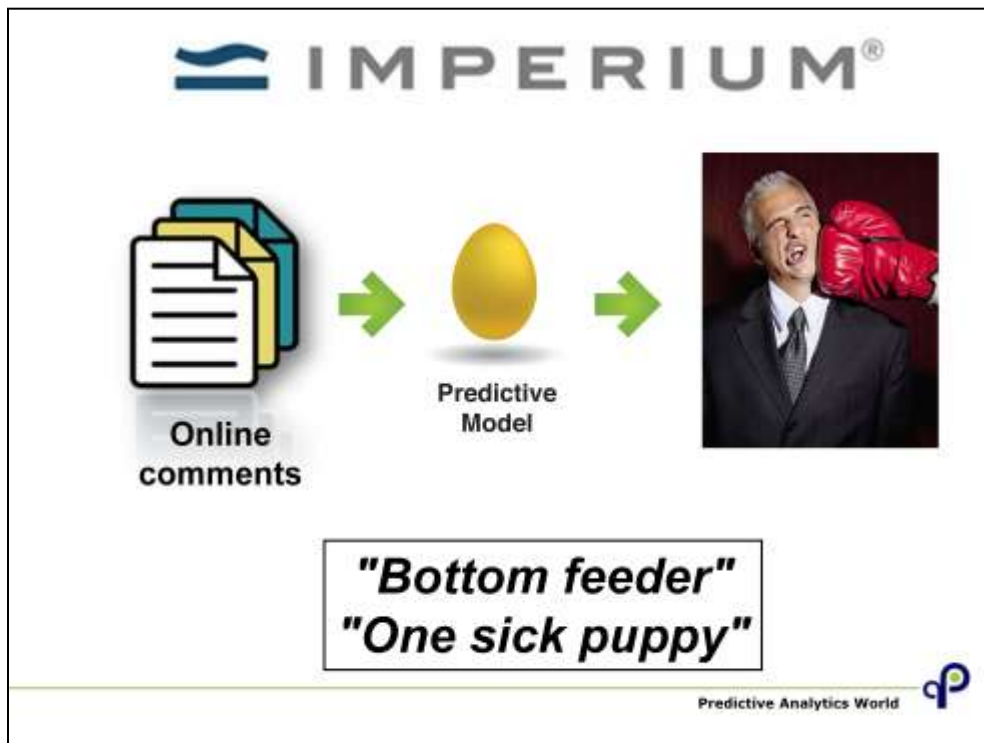
See also the paper S. Kim et al, “**Beyond Sentiment: The Manifold of Human Emotions**”

<http://arxiv.org/abs/1202.1568>

<http://arxiv.org/pdf/1202.1568v2.pdf>

Also see analysis of news broadcast bias by textual analysis of closed caption text:

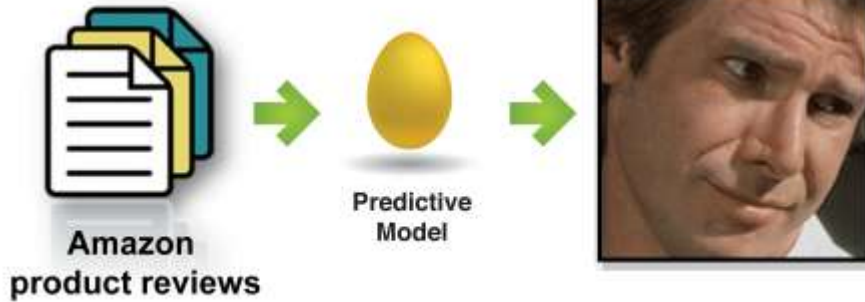
<http://arxiv.org/abs/1307.4879v1>




The data integrity company Imperium sponsored a competition to identify insults within online comments and blogs such as "bottom feeder" and "one sick puppy."

Kaggle, "Predict Whether a Comment Posted during a Public Discussion Is Considered Insulting to One of the Participants," Competition, August 7, 2012. www.kaggle.com/c/detecting-insults-in-social-commentary.

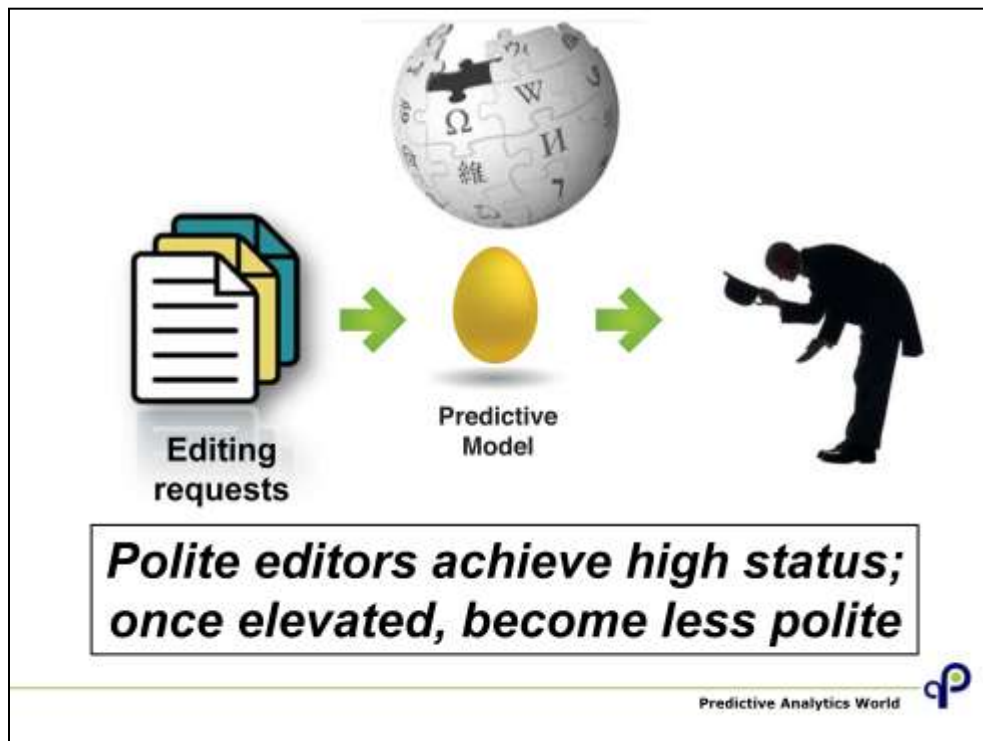
amazon.com



**Research model IDs 83%.
"Trees died for this book?"**

Predictive Analytics World 

Hebrew University identified 83 percent of sarcastic Amazon product reviews (e.g., "Trees died for this book?").



“Polite Wikipedia editors are more likely to achieve high status through elections, but, once elevated, they become less polite.”

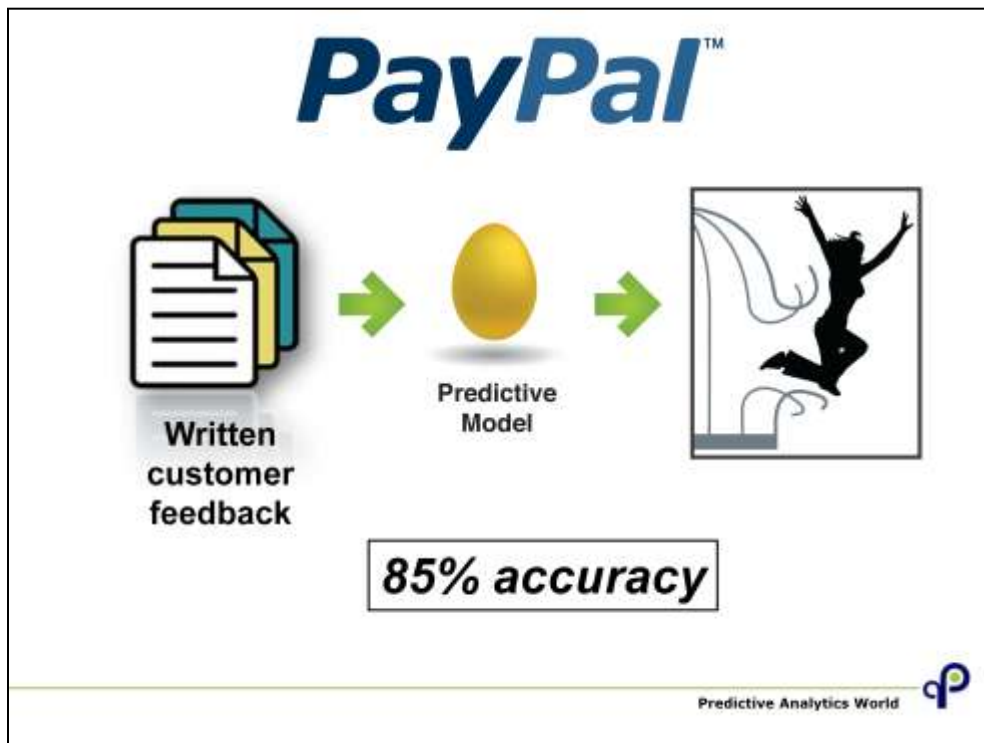
A Computational Approach to Politeness with Application to Social Factors, by Danescu-Niculescu-Mizil et al:

<http://arxiv.org/abs/1306.6078>

<http://www.mpi-sws.org/~cristian/Politeness.html>

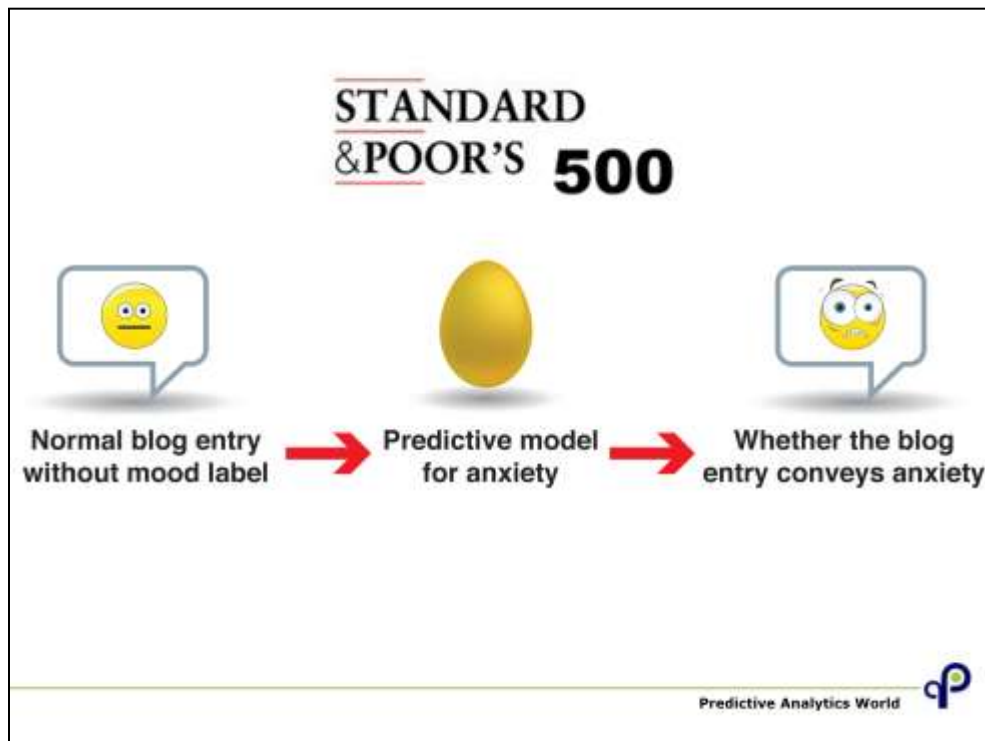
See also, determination of *controversial* Wikipedia articles:

<http://arxiv.org/abs/1305.5566>

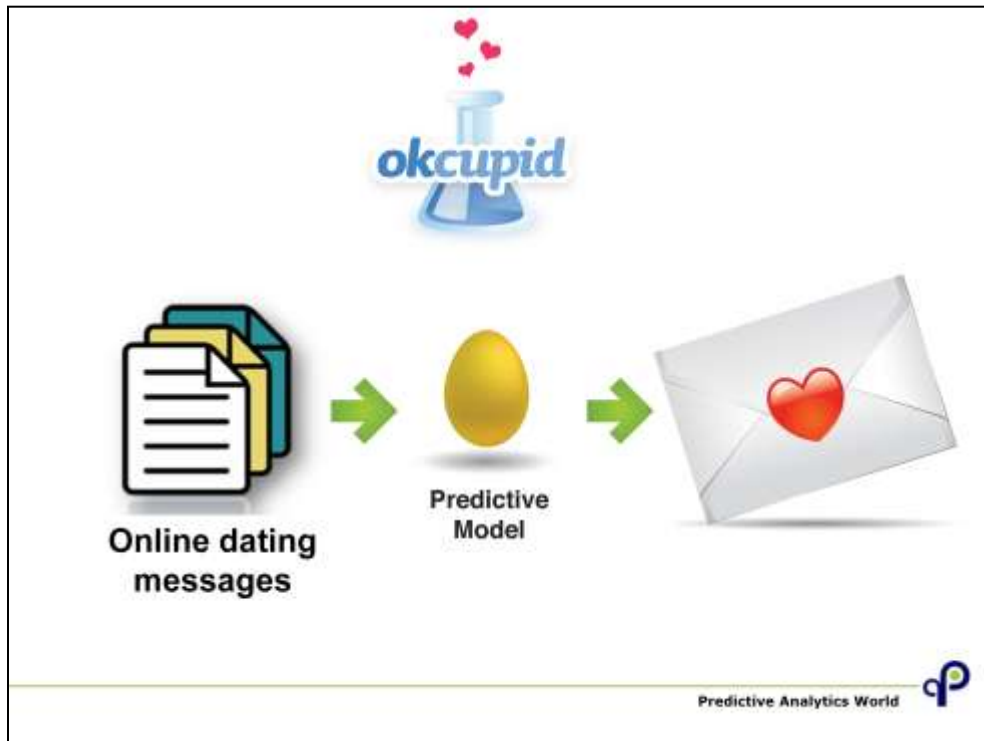


PayPal identifies from written feedback customers who intend to leave (aka churn or defect) with 85 percent accuracy.

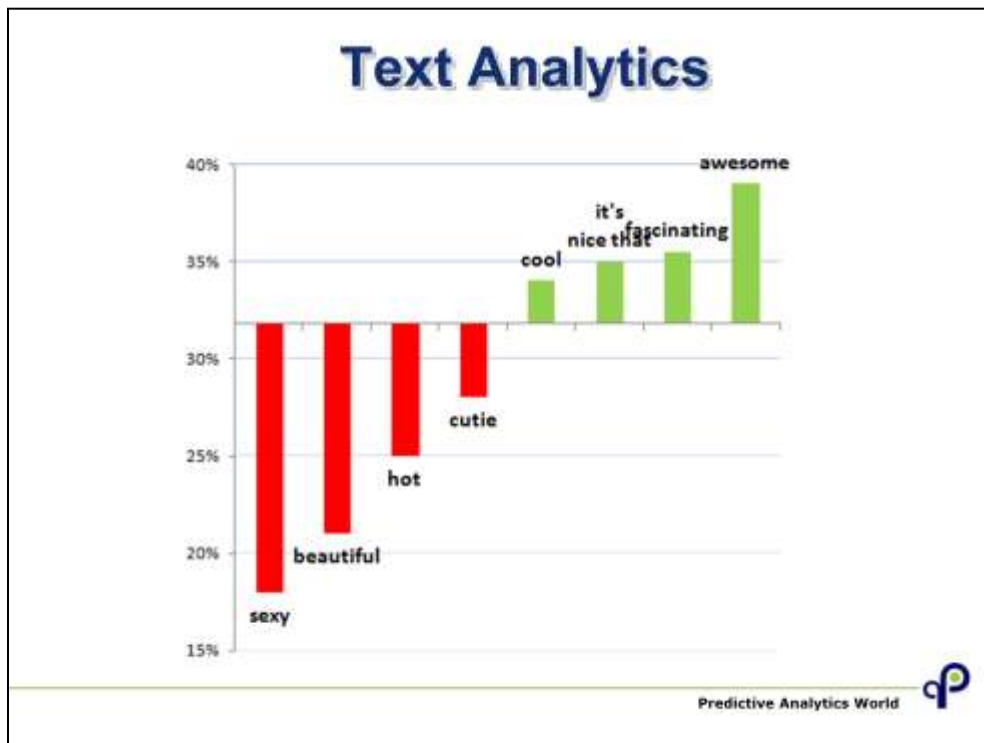
www.textanalyticsworld.com/newyork/2011/agenda/full-agenda#day1-gold



Although a rough approximation, these anxiety-predicting models promise to deliver insight about mass mood. While they manage to detect only 28 percent to 32 percent of each day's anxious blogs, this detection rate in fact serves quite well. What's important is measuring the relative prevalence of anxious blog entries on any given day. So, for example, if the number of anxious posts doubles from one day to the next, identifying just 30 percent of anxious posts each day will still reveal the overall increase in anxiety, even though many anxious posts are missed. Among blog entries labeled as anxious, the models are quite precise, mistaking not anxious entries as anxious only 3 percent to 6 percent of the time.

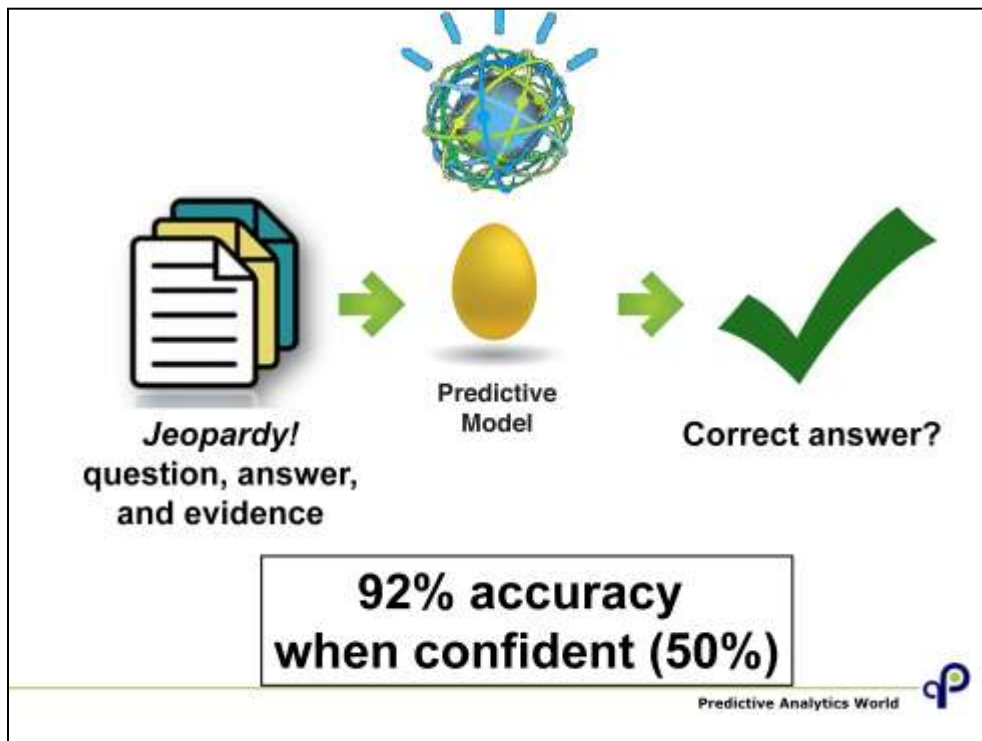


Hebrew University identified 83 percent of sarcastic Amazon product reviews (e.g., "Trees died for this book?").



Text (webpages, blogs, journals, messages [as in this slide], etc.) is a goldmine of human knowledge.

Chart from <http://blog.okcupid.com/>



IBM's Watson computer predicts, for an individual *Jeopardy!* question and candidate answer, whether it is the correct answer. For more information, see Chapter 6 of *Predictive Analytics* (www.thepredictionbook.com)

Four Ways Predictive Analytics Leverages Social Media

1. Predict behavior from who you know
2. Predict behavior from social activity
3. Predict what you like
4. **Predict who you know and your influence**



facebook

LinkedIn



Characteristics
of an Individual



Predictive
Model



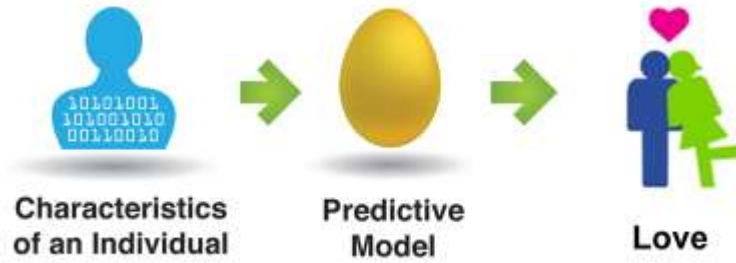
Friendships

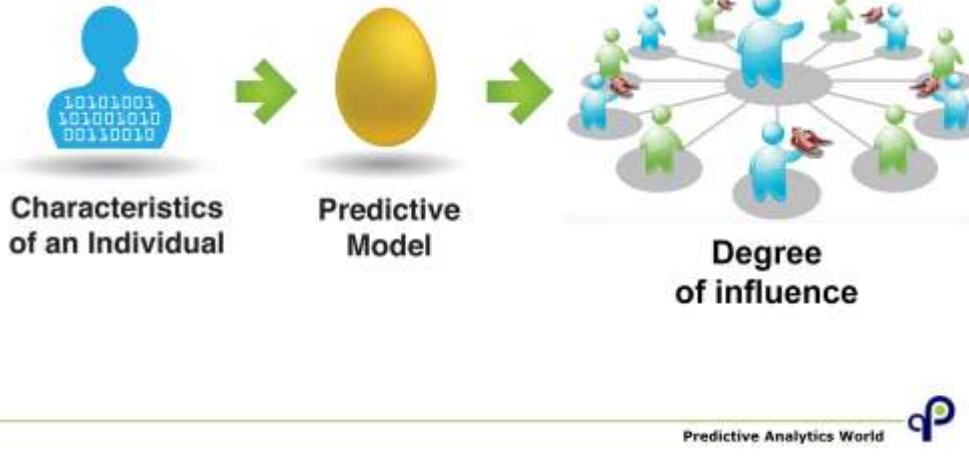
LinkedIn: this is
"the most important data product we built."

Predictive Analytics World

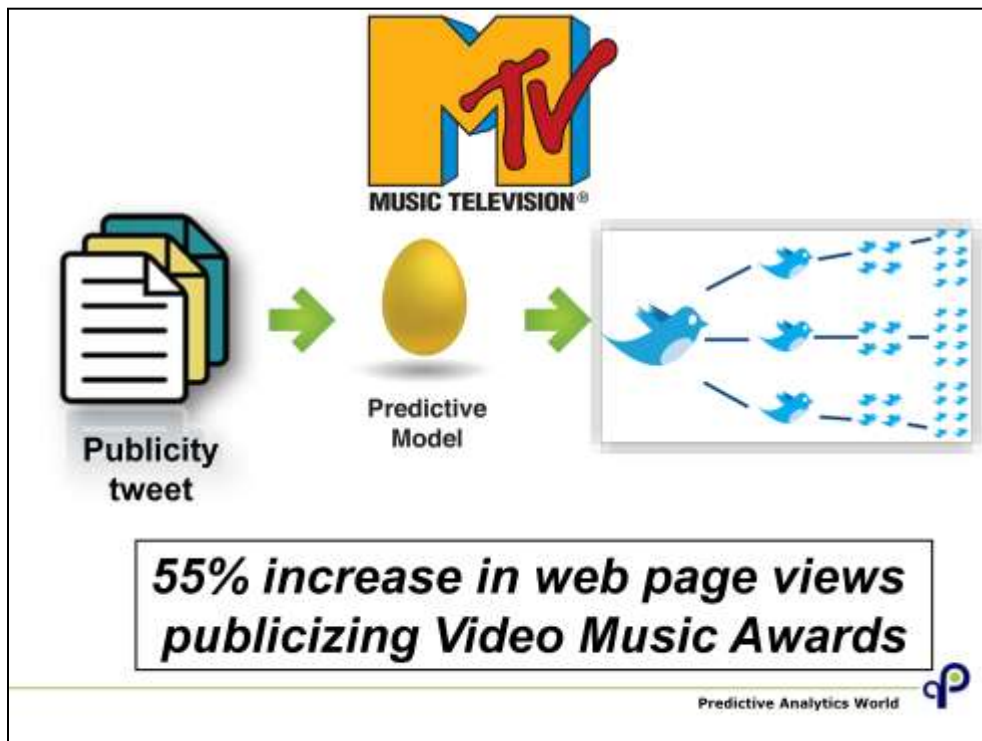


match.com





Microsoft works to predict which people are influential in a social network



See also, Predicting the Number of Likes on a Facebook Status With Statistical Keyword Analysis:

<http://minimaxir.com/2013/06/big-social-data/>

See also, "What Makes online Content Viral?" by Jonah Berger and Katherine L. Milkman, Journal of Marketing Research, American Marketing Association, ISSN: 0022-2437 (print), 1547-7193 (electronic)

Wrap-Up

- 1) Driving treatment selection with PA
- 2) Four ways PA leverages social media





Predictive Analytics World Conference

San Francisco • Boston
Chicago • Toronto
London • Berlin
Washington DC (Government)

Bigger wins!
Strengthen the business impact
delivered by predictive analytics.

www.pawcon.com

"Predictive Analytics World was probably the best analytics conference I have attended... turned into my new must-go-to conference."

Dennis Mortensen
Director of Data Insights
Yahoo!

PAW March 2012 drew 500 attendees. PAW has included case studies from: AT&T, Aflac, Amazon.com, Amway, Anheiser-Busch, BBC, Barclays, Blue Cross, BoA, Citibank, Fidelity, GE Capital, Google, HP, HSBC, Infinity, Intuit, LinkedIn, Macy's, Match.com, MetLife, Microsoft, NRA, Netflix, Orbitz, Overstock.com, PayPal, Reed Elsevier, Target, The Economist, US Bank, US Bank, Wells Fargo, YMCA, Yahoo! and many more.

91% of attendees rated
the instructor
Excellent or Very Good

P PREDICTION **IMPACT**

Predictive Analytics for Business and Marketing

Online training program:
Predictive Analytics Applied

Available on-demand, at your convenience

"Thank you for the perfect
introduction to predictive analytics.
It's just the right blend of a business
and technical angle I was looking for."

—Emiliano Pasqualetti,
CEO, DomainsBot Inc.



businessprediction.com

**To Learn More:
The Predictive Analytics Guide**

www.pawcon.com/guide

Predictive Analytics World

