



# Forging a Successful Path to Machine Learning Deployment

CAS Annual Meeting  
November 2021

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

*To continue your learning beyond this keynote:*

## **Machine Learning Leadership and Practice – End-to-End Mastery**

This end-to-end, three-course will empower you to launch machine learning. Accessible to business-level learners and yet vital to techies as well, it covers both the state-of-the-art techniques and the business-side best practices.

<http://www.MachineLearning.courses>

Also, for more information/citations regarding the examples in this presentation, see the Notes, freely-accessible online, for the book "Predictive Analytics" by Eric Siegel (<http://www.thepredictionbook.com>). Most of the various examples shown are covered in the book (some only briefly, within the book's Central Tables of 182 mini-case studies, so not necessarily with more detail there than in this presentation). So, for greater detail about each case study named, see its reference/citation - search by organization name within the book's Notes PDF, available online at <http://www.PredictiveNotes.com>.

**data** **business**  
**machine** **prediction**  
**learning**







# ***predictive analytics***



**boost sales**

**cut costs**

**combat risk**

**prevent fraud**

**fortify healthcare**

**streamline manufacturing**

**conquer spam**

**bolster insurance**



**But isn't prediction impossible?**

**predictive analytics**



# Agenda

1. What is machine learning / predictive analytics?
2. How it delivers value
  - Business applications
  - Insurance
3. How it works
  - *The Prediction Effect*
  - *The Data Effect*
4. Forging a path to deployment



*the* **PREMIER**  
**MACHINE LEARNING**  
EVENT

**JUNE 19-24, 2022**

**MACHINE  
LEARNING  
WEEK**

**CAESARS PALACE  
LAS VEGAS**



[www.pawcon.com](http://www.pawcon.com)

<https://www.predictiveanalyticsworld.com/>

ONLINE COURSE SERIES

## Machine Learning Leadership and Practice – *End-to-End Mastery*

Learn the state-of-the-art  
techniques and the business-  
side best practices.

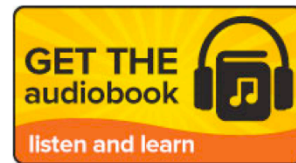
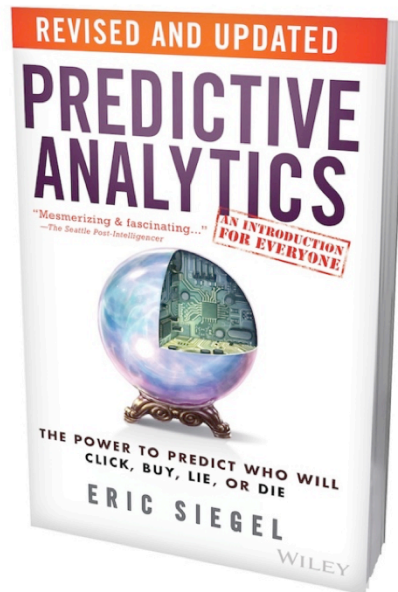


[www.MachineLearning.courses](http://www.MachineLearning.courses)

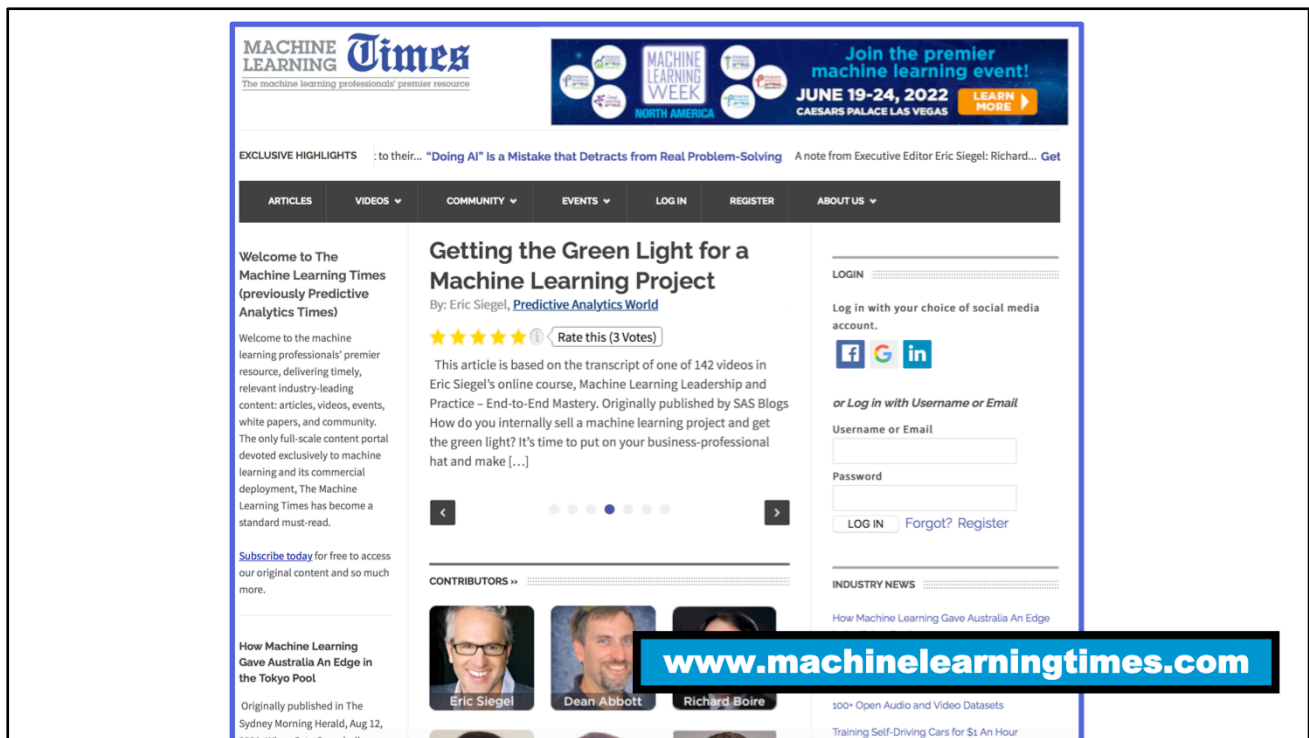
### **Machine Learning Leadership and Practice – End-to-End Mastery**

This end-to-end, three-course will empower you to launch machine learning. Accessible to business-level learners and yet vital to techies as well, it covers both the state-of-the-art techniques and the business-side best practices.

<http://www.MachineLearning.courses>



[www.ThePredictionBook.com](http://www.ThePredictionBook.com)



The Machine Learning Times has become the destination for the constantly evolving machine learning community.

<http://www.machinelearningtimes.com>

This site is the machine learning professionals' premier resource, delivering timely, relevant industry-leading content: articles, videos, events, white papers, and community. The only full-scale content portal devoted exclusively to predictive analytics and its commercial deployment, the Predictive Analytics Times has become a standard must-read.





**@predictanalytic**



**Before:**



**After:**



I am an individual patient, and an individual insurance policyholder. Risk effects all parties involved.

## Knee Walking



Eric Siegel @predictanalytic

[www.MachineLearning.courses](http://www.MachineLearning.courses)

Predictive Analytics World



ACL replacement surgery choice of graft source influences the risk of long term knee pain when “knee walking”.



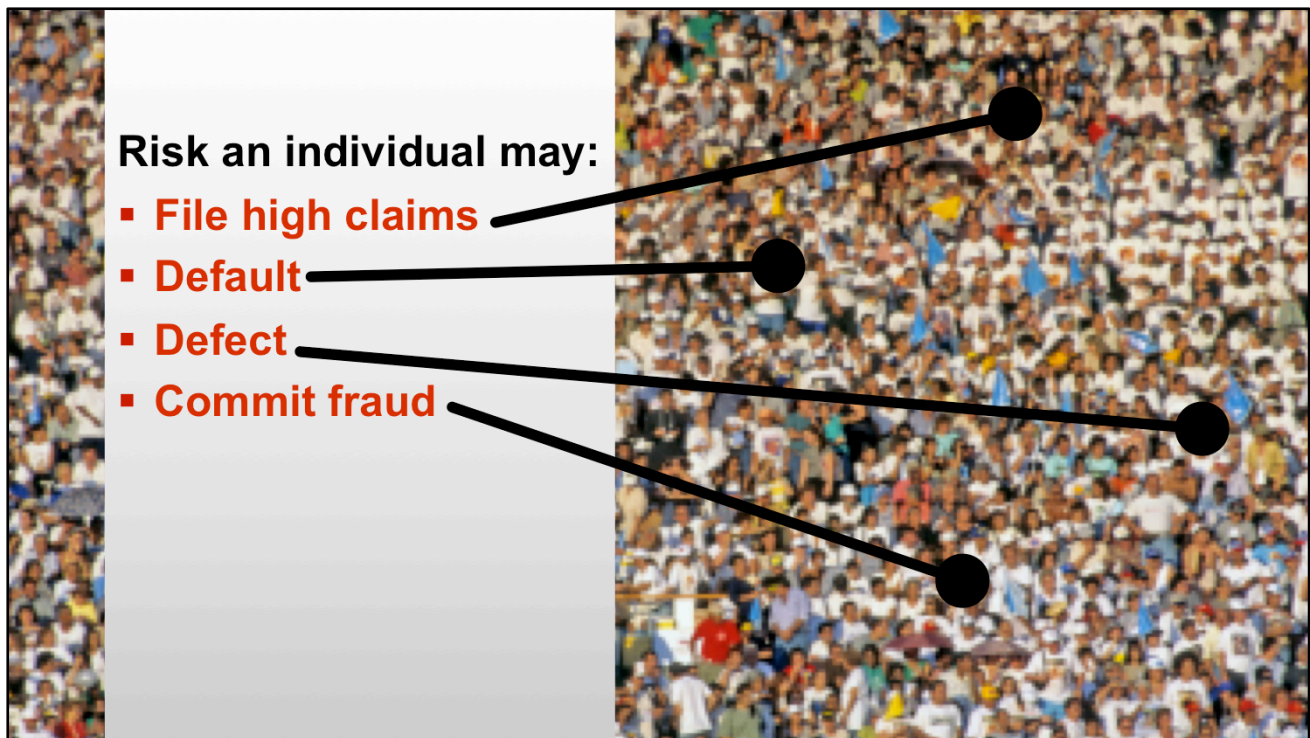
Insured “office workers”

“*Insurance is nothing but management of information. It is pooling of risk, and whoever can manipulate information the best has a significant competitive advantage.*”

**Eric Webster**  
VP Marketing, State Farm



"Discussion with State Farm's Eric Webster: Insurance and Data Mining," Gregory Piatetsky, Ph.D., KD Nuggets. <http://www.kdnuggets.com/news/2009/n08/3i.html>



**Risk an individual may:**

- **File high claims**
- **Default**
- **Defect**
- **Commit fraud**

“  
*No certified, regulated profession like the actuarial practice exists outside of what is strictly considered insurance.*”

**Douglas Hubbard**  
*The Failure of Risk Management*



Also: <http://www.prnewswire.com/news-releases/deloitte-poll-37-percent-of-companies-report-having-no-consistent-risk-data-mining-reporting-and-analysis-116644679.html>

# Data Scientist:

## *The Sexiest Job of the 21st Century*

**Meet the people who  
can coax treasure out of  
messy, unstructured data.**  
by Thomas H. Davenport  
and D.J. Patil



<https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/>

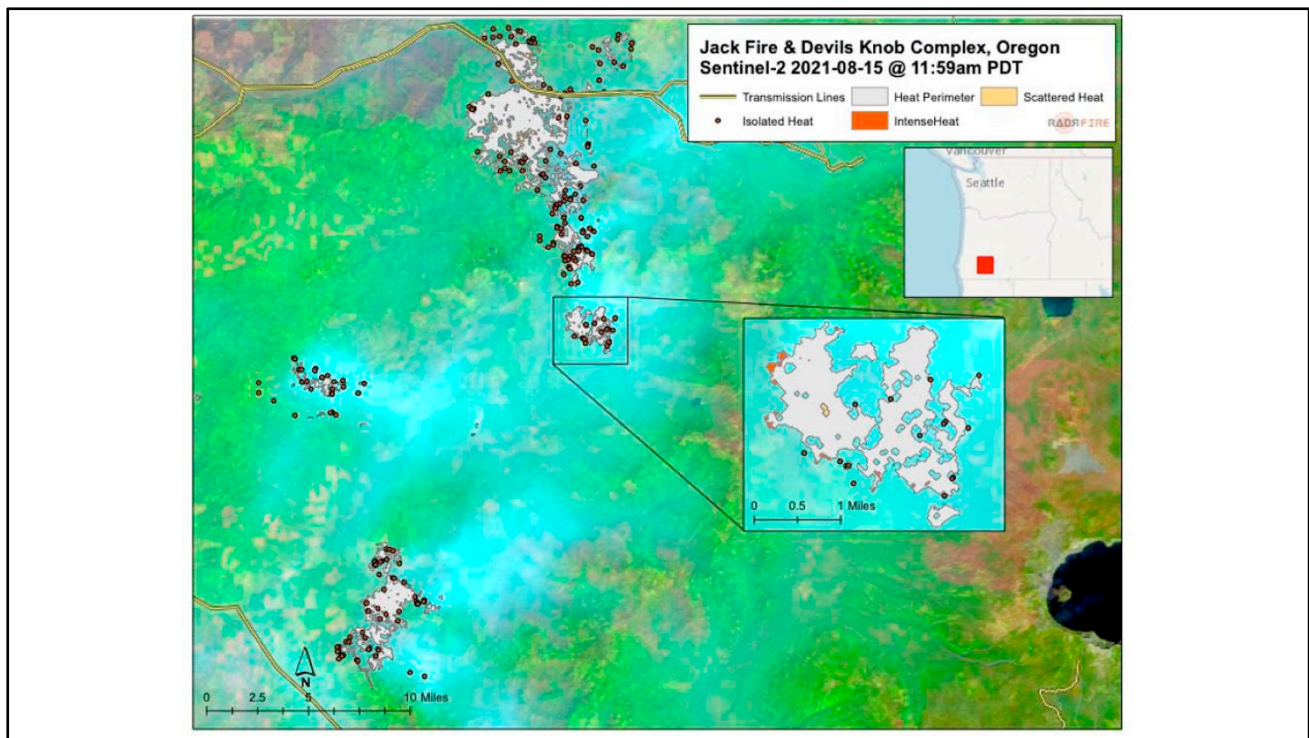












Machine learning to detect wildfires

<https://www.forbes.com/sites/jamesconca/2021/08/31/capturing-wildfires-from-space--pnnls-radrfire/?sh=261c85a57d6c>

<https://www.scientificamerican.com/article/ai-could-spot-wildfires-faster-than-humans/>

<https://www.wsj.com/articles/california-firefighters-tap-ai-for-an-edge-in-battling-wildfires-11601544600>



**Con Edison:** Predicts dangerous manhole explosions and fires in New York City, identifying a 2 percent of manholes that have a 5.5 times greater than average risk of an incident.



Predictive analytics educational/infotainment geek rap video:

<http://www.PredictThis.org>



# Machine learning:



**(data)**

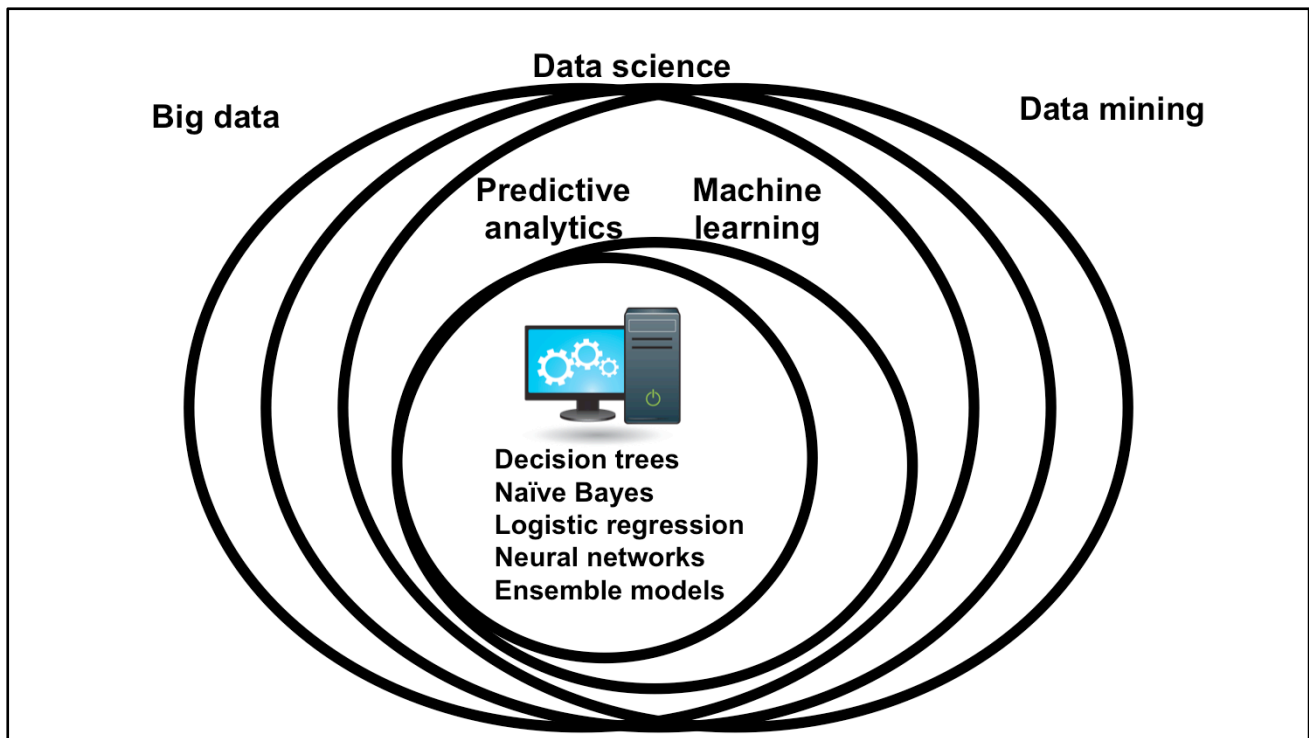
*Technology that learns from experience to predict the outcome or behavior of each customer, patient, business, vehicle, image, piece of equipment, or other individual unit  
... in order to drive better decisions.*

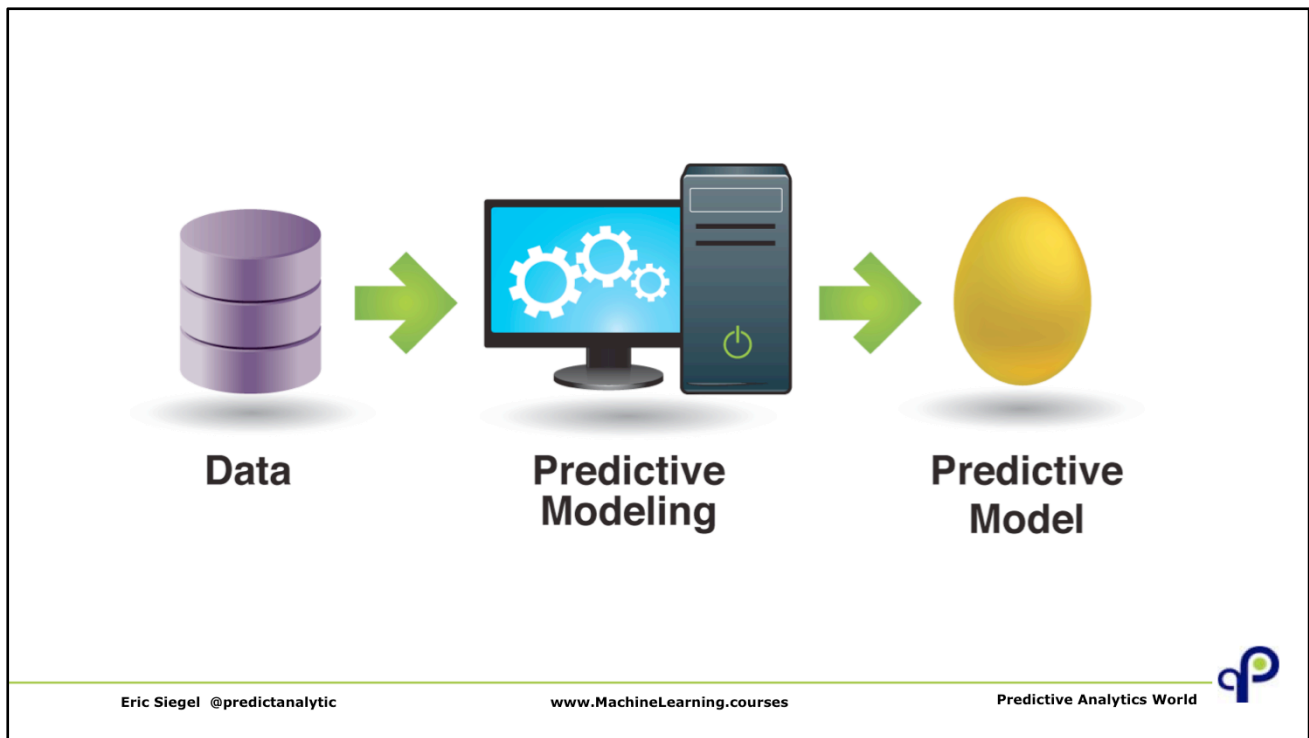


This talk is about machine learning in the above practical, applied sense.

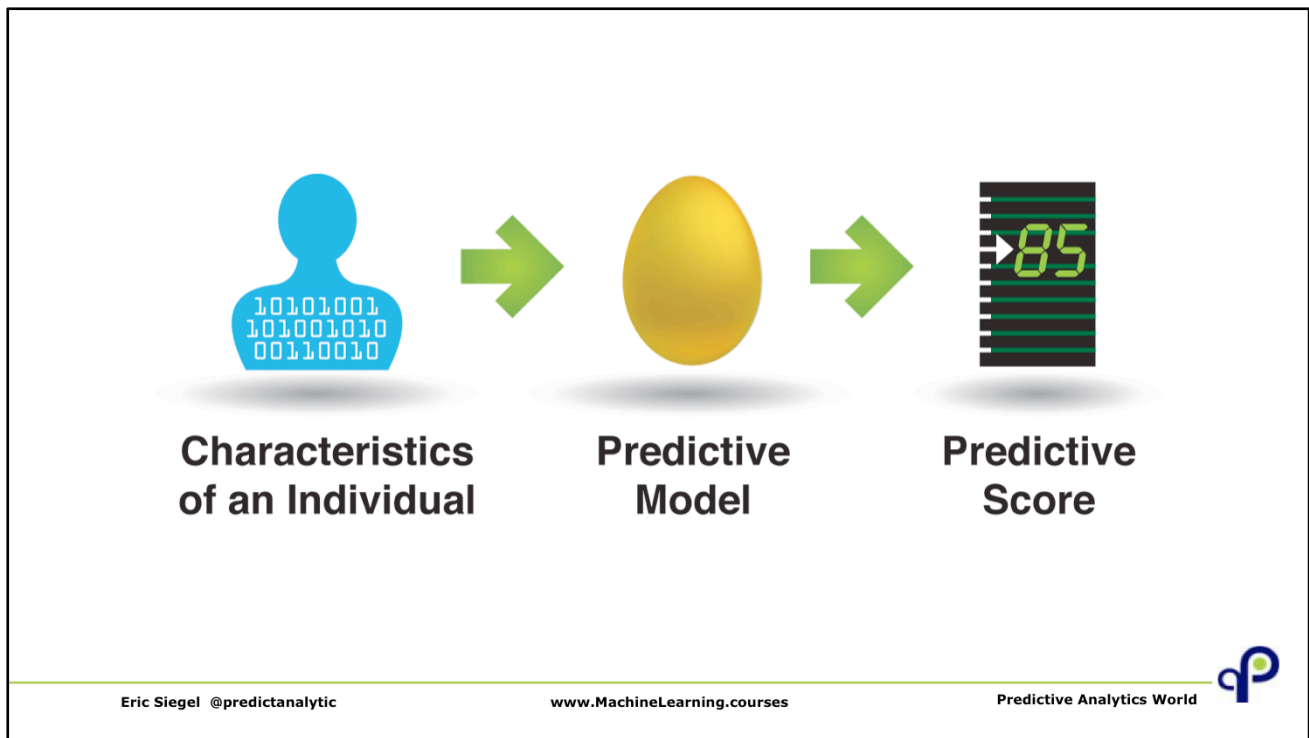
A.k.a. *predictive analytics*.



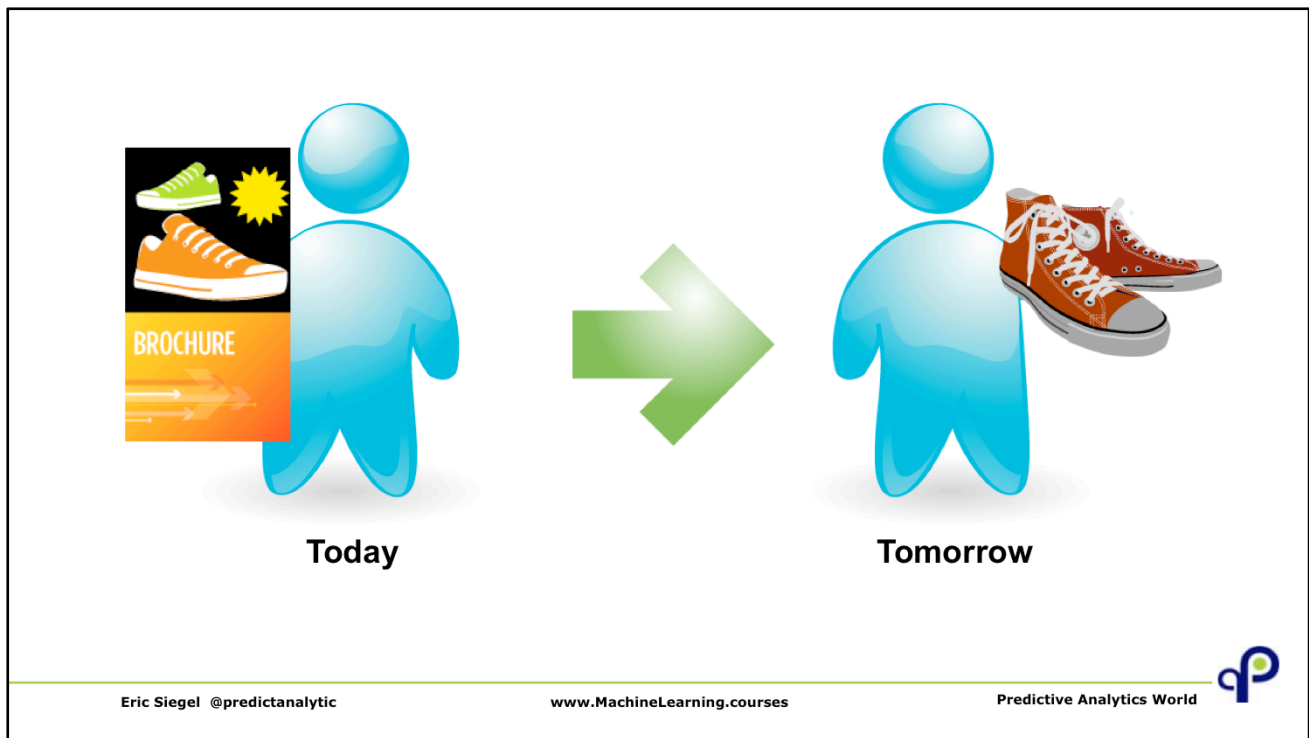




Predictive modeling learns from data in order to generate a predictive model. For details on how this works, see Chapter 4 of the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" (<http://www.thepredictionbook.com>).



A predictive model generates a predictive score for an individual. For details on how this works, see Chapters 1 and 4 of the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" (<http://www.thepredictionbook.com>).

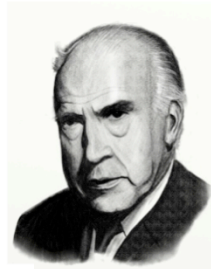


Marketing targets an individual predicted as likely to buy. For details on how this works see the Introduction and Chapter 1 of the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" (<http://www.thepredictionbook.com>).

# The Challenge of Prediction

*Prediction is very difficult, especially if it's about the future.*

- Niels Bohr



*How come you never see a headline like "Psychic Wins Lottery"?*

- Jay Leno



Is prediction an audacious goal? Isn't prediction impossible? For details on how why predictive analytics predicts well enough, see the Introduction and Chapter 1 of the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" (free to access as a PDF on the "Excerpts" page of <http://www.thepredictionbook.com>).

# The Accuracy Fallacy

**NEWSWEEK:** *AI Can Tell If You're Gay: Artificial Intelligence Predicts Sexuality from One Photo with Startling Accuracy*

**THE SPECTATOR:** *Linguistic Analysis Can Accurately Predict Psychosis*

**THE DAILY MAIL:** *AI-Powered Scans Can Identify People at Risk of a Fatal Heart Attack Almost a Decade in Advance*

**THE NEXT WEB:** *This Scary AI Has Learned How to Pick Out Criminals by Their Faces*

Eric Siegel @predictanalytic

[www.MachineLearning.courses](http://www.MachineLearning.courses)

Predictive Analytics World



For more on the accuracy fallacy, see: "The Media's Coverage of AI is Bogus", by Eric Siegel, Scientific American, November 2019. <https://blogs.scientificamerican.com/observations/the-medias-coverage-of-ai-is-bogus/>

[Subscribe](#)
[Latest Issues](#)

**SCIENTIFIC AMERICAN**

[Cart 0](#)
[Sign In](#)
[Newsletters](#)

---

# The Accuracy Fallacy

Claims that machine learning can predict sexuality, psychosis, and more are greatly overblown

By Eric Siegel on November 20, 2019



Credit: Valery Brozhinsky Getty Images

Headlines about machine learning promise godlike predictive power. Here are four examples:

- *Newsweek*: "AI Can Tell If You're Gay: Artificial Intelligence

**READ THIS NEXT**

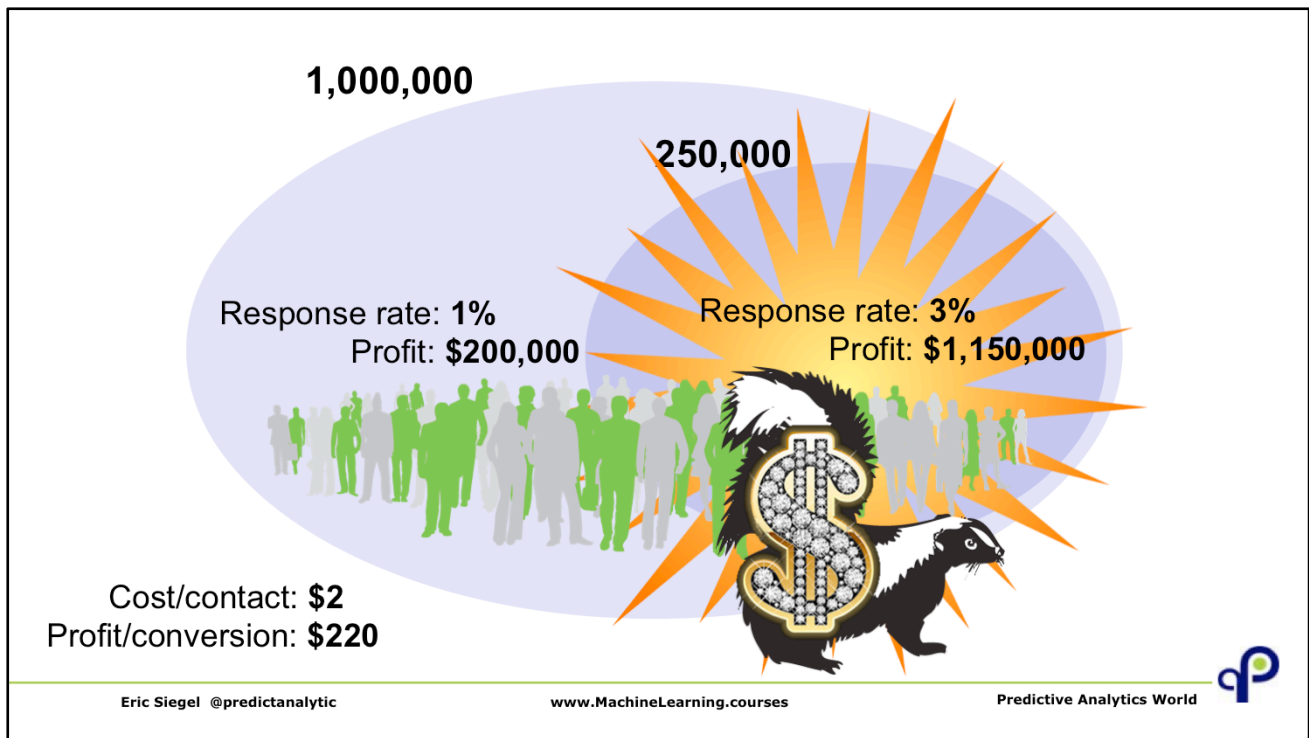
**BIOLOGY**  
Guardians of the Microbial Galaxy  
6 hours ago — Janani Hantharani | Opinion

**PUBLIC HEALTH**  
Data and Technology Can Help Us Make Progress on COVID Inequities  
March 27, 2021 — Daniel E. Dawes and Karen DeSalvo | Opinion

**PUBLIC HEALTH**  
Coronavirus News Roundup, March 20 - March 26  
March 26, 2021 — Robin Lloyd | Opinion

**CLIMATE**  
Solar Geoengineering Should be Investigated, Scientists Say

<https://blogs.scientificamerican.com/observations/the-medias-coverage-of-ai-is-bogus/>



A crummy predictive model delivers big value. It's like a skunk with bling.

Simple arithmetic shows the bottom line profit of direct mail, both in general and then improved by predictively targeting (and only contacting 25% of the list). The less simple part is how the predictive scores are generated for each individual in order to determine exactly who belongs in that 25%. For details on how this works, see Chapter 1 of the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" (<http://www.thepredictionbook.com>).



## The Prediction Effect:

*A little prediction goes a long way.*



Put another way, predicting better than guessing is often sufficient to generate great value by rendering operations more efficient and effective. For details on how this works, see the Introduction and Chapter 1 of the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" (available for free on the Excerpts page of <http://www.thepredictionbook.com>).



#### FINANCIAL SERVICES

*Lowered direct mail costs 20%  
Increased response rate 3.1%  
600% ROI*



#### RETAIL

*Improved direct mail  
targeting by 15-20%*



#### FINANCIAL SERVICES

*Reduced mailing costs by  
\$12 million*

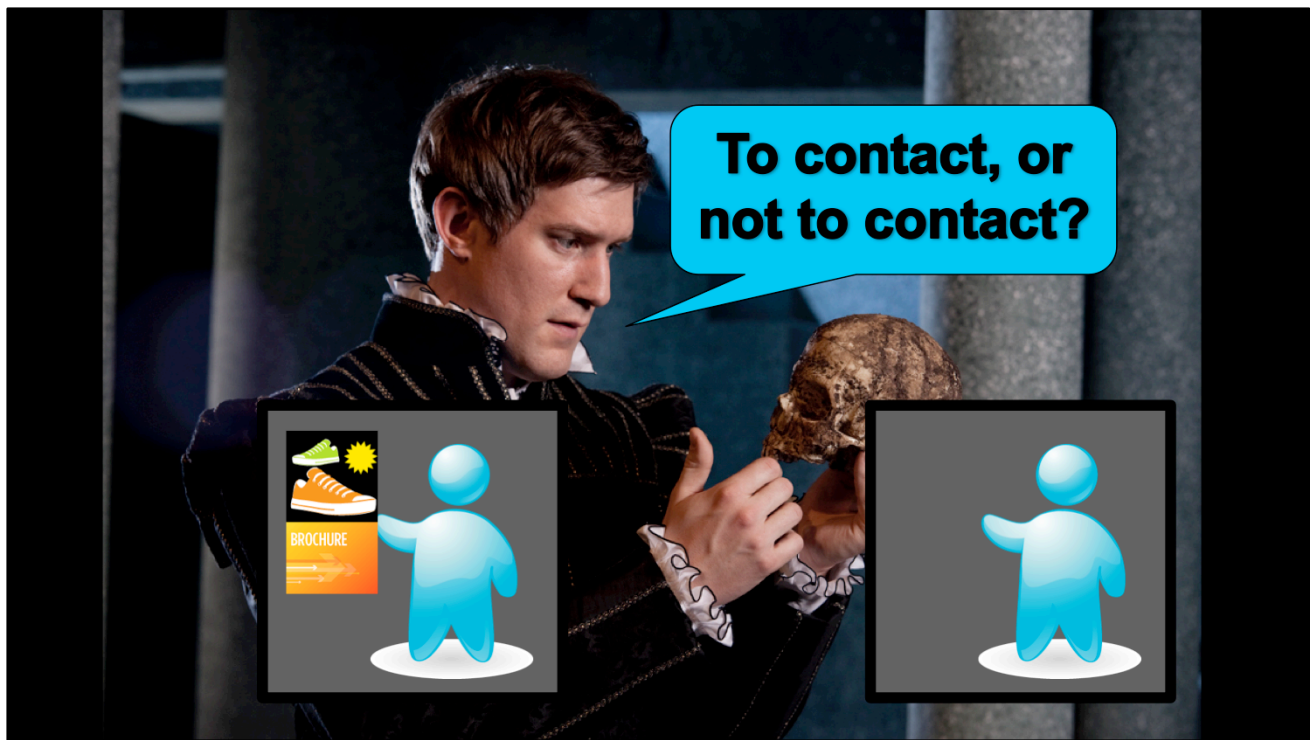
...and many more, such as Cox Communications, FedEx, Sprint, etc. - see the book "Predictive Analytics" ([www.thepredictionbook.com](http://www.thepredictionbook.com)) for many case studies, including a central compendium of 147 mini-case studies, of which 37 are examples in marketing applications of predictive analytics.

Reference for most examples/case studies in this presentation are in the Notes PDF for Eric Siegel's book, "Predictive Analytics." For each example's reference/citation, search by organization name within the book's Notes PDF, available at [www.PredictiveNotes.com](http://www.PredictiveNotes.com)

PREMIER Bankcard also lowered delinquency to increase net by over \$10 million

More information about First Tennessee Bank and other case studies are available at <http://tinyurl.com/PAExamples>

Dan Marks, First Tennessee Bank, "First Tennessee Bank: Analytics Drives Higher ROI from Marketing Programs," IBM.com, March 9, 2011. [www.ibm.com/smarterplanet/us/en/leadership/firsttenbank/assets/pdf/IBM-firstTennBank.pdf](http://www.ibm.com/smarterplanet/us/en/leadership/firsttenbank/assets/pdf/IBM-firstTennBank.pdf)



## ***Antidote to Information Overload:*** **PREDICTION**

-  search results
-  orders news feed
-  predicts viral posts
- Email filters spam
- Retailers diminish "junk mail"
- Movies 
- Music 
- Books 
- Flights 
- Rentals 
- Dates 

Eric Siegel @predictanalytic

[www.MachineLearning.courses](http://www.MachineLearning.courses)

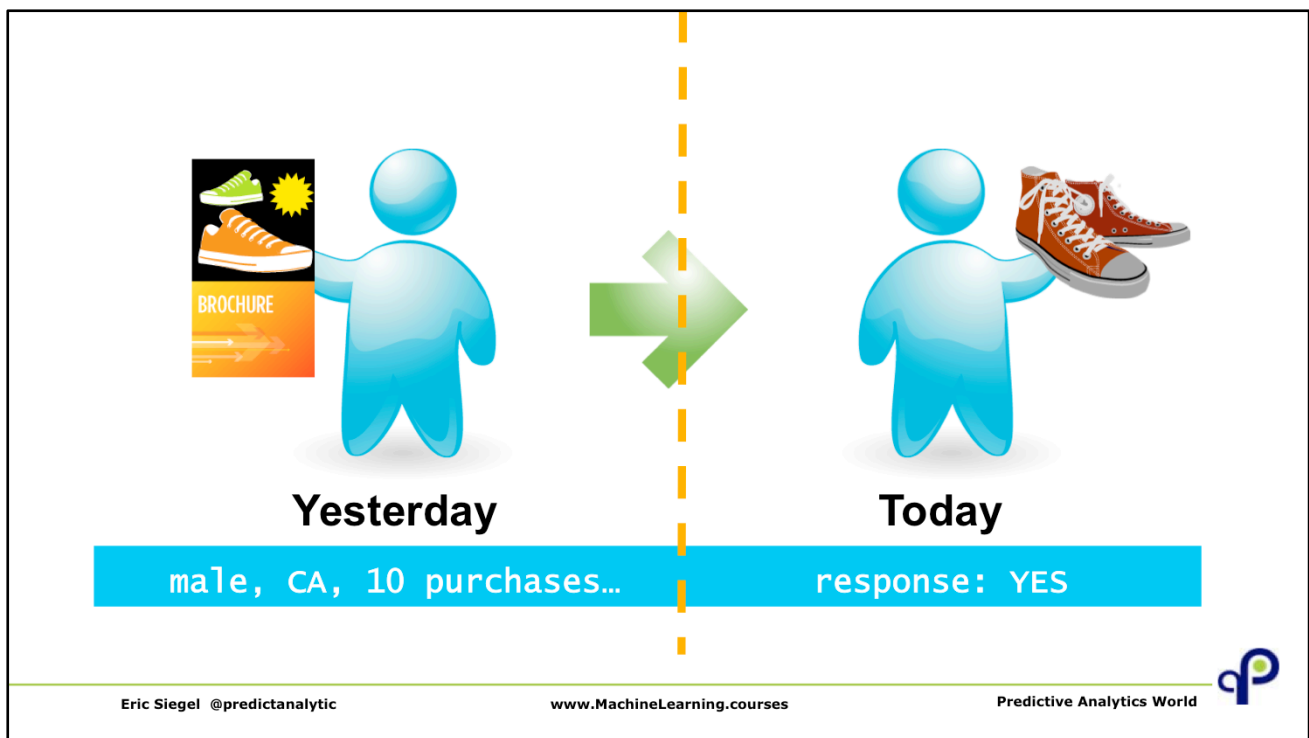
Predictive Analytics World



More examples:

- person to tag in a photo (Facebook)
- ad that interests you (also Facebook)
- people to "friend" (Facebook and LinkedIn)
- least-congested driving route (Waze and IBM)
- optimal shipping routes (for UPS drivers).





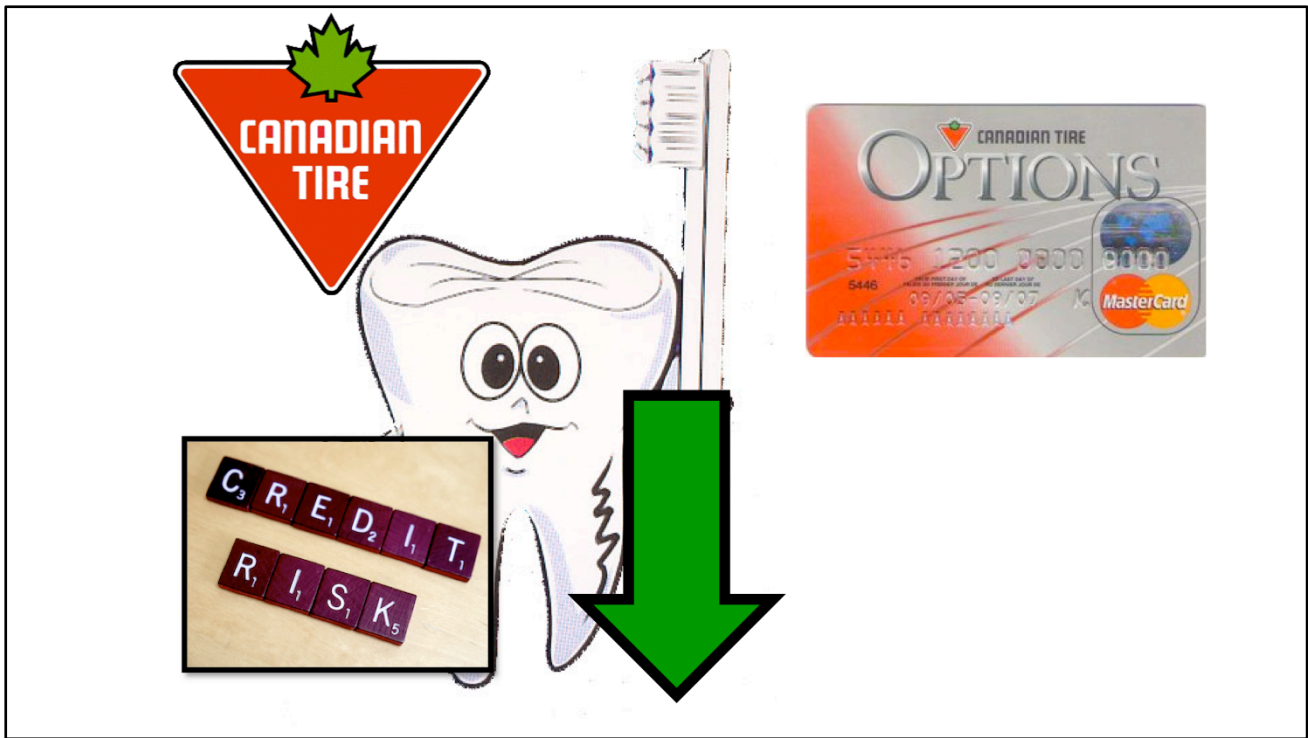
Each row of *training data* corresponds to one individual – first the individual’s facts and figures are listed (predictor variables, aka independent variables), and then the target variable (aka dependent variable) – ie, the thing you’re trying to predict – is listed.

A table of such rows composes the training data, on which predictive modeling operates.



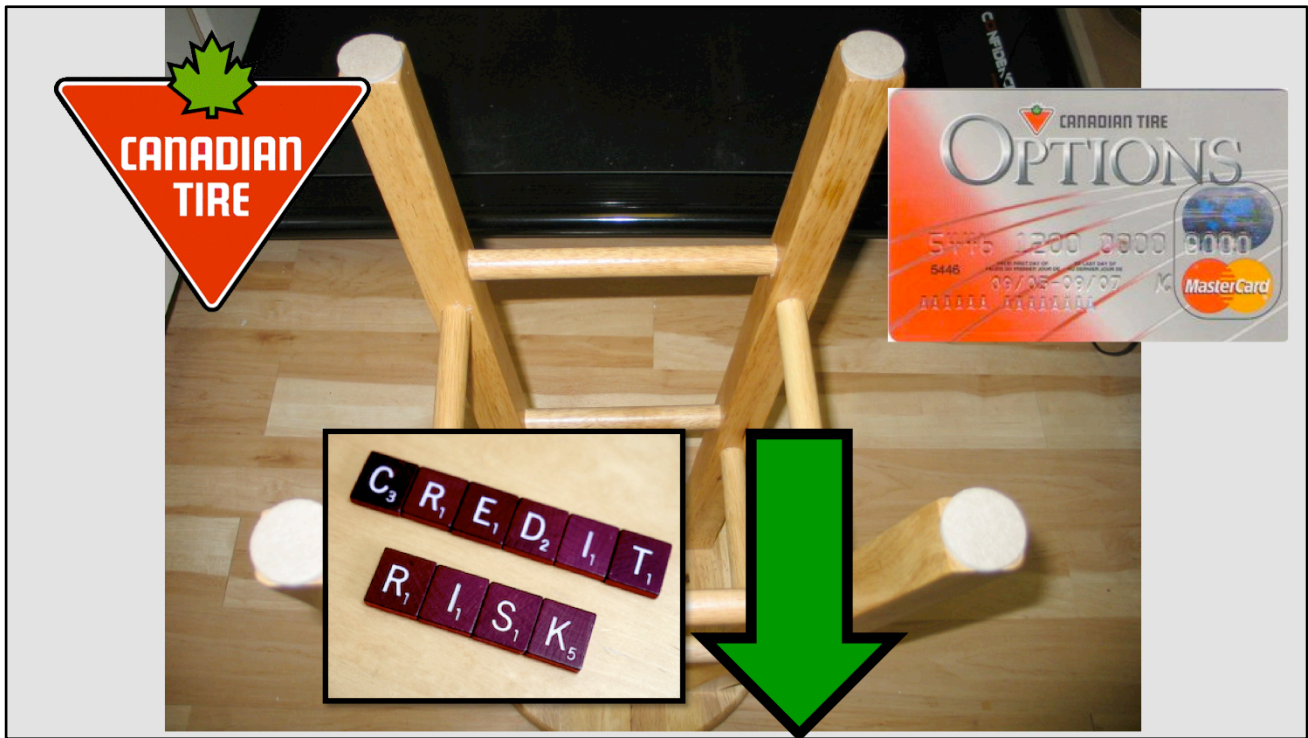


Canadian Tire examples, from "What Does Your Credit-Card Company Know About You?" New York Times, May 12, 2009. <http://www.nytimes.com/2009/05/17/magazine/17credit-t.html>

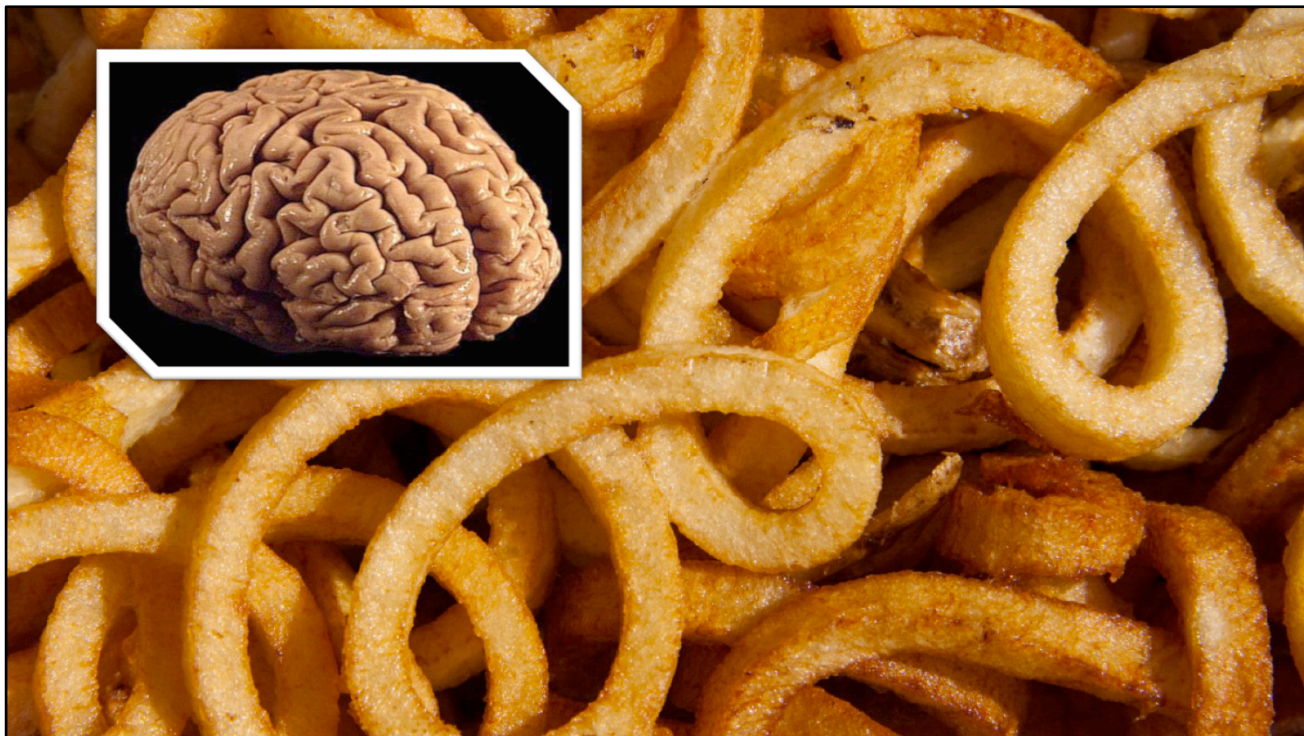


Canadian Tire examples, from "What Does Your Credit-Card Company Know About You?" New York Times, May 12, 2009. <http://www.nytimes.com/2009/05/17/magazine/17credit-t.html>



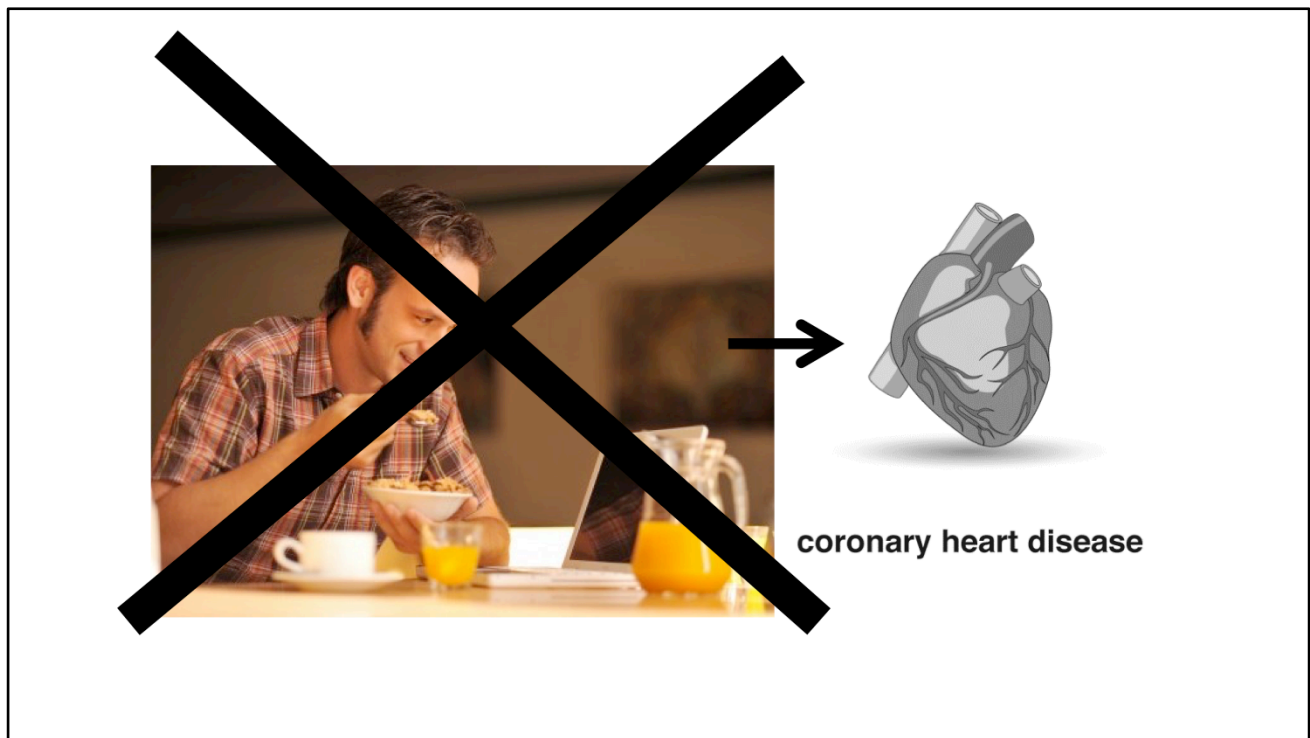


Canadian Tire examples, from "What Does Your Credit-Card Company Know About You?" New York Times, May 12, 2009. <http://www.nytimes.com/2009/05/17/magazine/17credit-t.html>



People who “like” curly fries on Facebook are “more intelligent.”

Reference for most examples in this presentation are in the Notes PDF for Eric Siegel's book, "Predictive Analytics." For each example's reference/citation, search by organization name within the book's Notes PDF, available at [www.PredictiveNotes.com](http://www.PredictiveNotes.com)



**Men who skipped breakfast showed a 27% higher risk of coronary heart disease.**

"Prospective Study of Breakfast Eating and Incident Coronary Heart Disease in a Cohort of Male US Health Professionals," by Cahill et al.  
<http://circ.ahajournals.org/content/128/4/337.full.pdf>



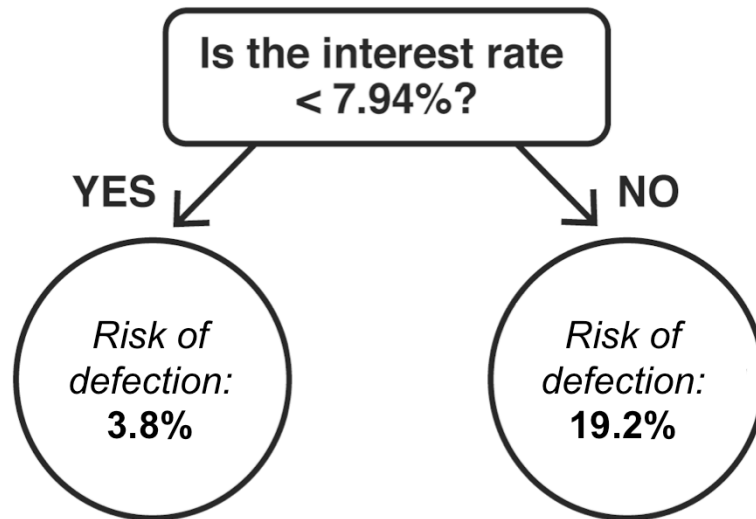
Neighborhoods of San Francisco that exhibit higher rates of certain crimes also exhibit higher demand for Uber rides.

# The Data Effect:

*Data is always predictive.*

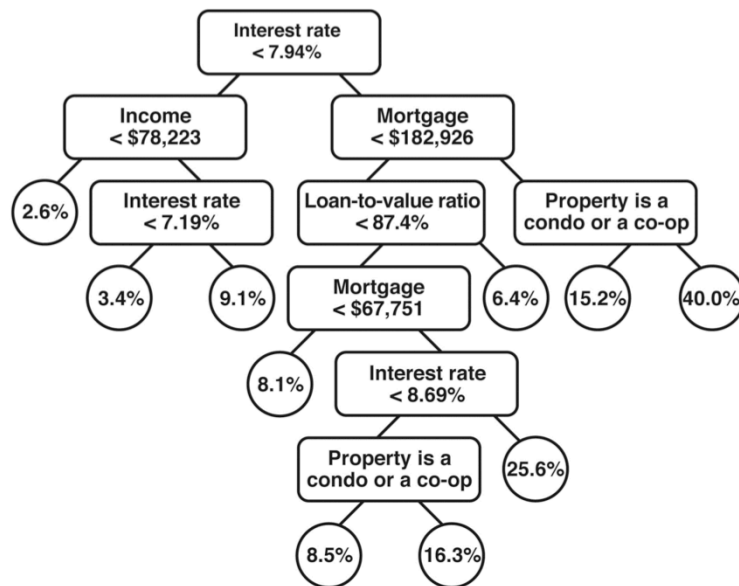


## Customer Attrition: Mortgages



Reference for most examples/case studies in this presentation are in the Notes PDF for Eric Siegel's book, "Predictive Analytics." For each example's reference/citation, search by organization name within the book's Notes PDF, available at [www.PredictiveNotes.com](http://www.PredictiveNotes.com)





“Insurance has always been about predictive analytics. What are actuarial tables, loss history analysis, and pricing/risk algorithms if not “predictive”?”

**Seth Earley**  
Earley & Associates



[http://uk.smartgridupdate.com/fc\\_fcbi1lz/lz.aspx?  
p1=056892S4006&CC=&p=1&clD=0&cValue=1](http://uk.smartgridupdate.com/fc_fcbi1lz/lz.aspx?p1=056892S4006&CC=&p=1&clD=0&cValue=1)



Insurance → ML

ML → Insurance ?



# Advantages of Machine Learning

**1. Data elements**

**2. Mathematical models**



## Predictive Analytics' Improvement:

***A leading international commercial lines insurance provider***

*Predictive models decreased the loss ratio by a half point, contributing to savings of almost \$50 million.*



*With predictive modeling, tripled the accuracy of predicting bodily injury liability - worth an estimated \$40 million annually.*



*Ascertains secondary medical conditions from workers' compensation claim notes; predictive of high-cost injuries.*

Leading international commercial lines insurance provider: This was from a case study at a conference presentation; however, the insurance company later rescinded authorization to be named in connection with this example. Here is a reference on the general principle that predictive modeling improves upon standard actuarial methods: Guillaume Briere-Giroux, FSA, MAAA, CFA, "Predictive Modeling Applied to Variable Annuity Lapse Rates," Predictive Modeling for Variable Annuities. [www.claudepenland.com/2011/02/12/predictivemodeling-applied-to-variable-annuity-lapse-rates/](http://www.claudepenland.com/2011/02/12/predictivemodeling-applied-to-variable-annuity-lapse-rates/).

Allstate: With a predictive modeling competition in 2012, tripled the accuracy of predicting bodily injury liability based solely on the characteristics of the insured vehicle. This could be worth an estimated \$40 million annually to the company -- Kaggle, Allstate, "Predicting Liability for Injury from Car Accidents," Competition, April 13, 2012. [www.kaggle.com/host/casestudies/allstate](http://www.kaggle.com/host/casestudies/allstate). Here is how the \$40 million estimate is derived. Allstate's 2010 annual report puts bodily injury claims at roughly \$1.6 billion. Industry insiders suggest 20 percent of bodily injury claims actuarially relate to vehicle. If we assume that a 1 percent improvement in error rate impacts claims costs by 0.25 percent, even just doubling the accuracy comes to a \$40 million savings. See also [www.iihs.org/research/hldi/fact\\_sheets/BodilyInjury\\_0911.pdf](http://www.iihs.org/research/hldi/fact_sheets/BodilyInjury_0911.pdf).

Accident Fund Insurance: Ascertains secondary medical conditions (such as obesity and diabetes) from written workers' compensation claim notes. These conditions are

# Fast-Tracking Claims



Increased the fast-tracking of claims by 1,100% with predictive analytics



# Fraud Detection



The National Insurance Crime Bureau says that **insurance criminals steal over \$30 billion annually.**

*This results in \$200 to \$300 of additional insurance premiums per U.S. household.*



Progressive Insurance Special Investigative Unit Report from the National Insurance Crime Bureau (NICB). [www.progressiveagent.com/claims/special-investigations-unit.aspx](http://www.progressiveagent.com/claims/special-investigations-unit.aspx).

This makes such fraud the second most costly white-collar crime in the United States - behind tax evasion.

## When Organizations Predict When You Will Die

A top-five U.S. health insurance company....  
... predicts the likelihood an elderly policyholder will  
pass away within 18 months...  
... in order to trigger end-of-life counseling, e.g.,  
regarding living wills and palliative care.

Eric Siegel @predictanalytic

[www.MachineLearning.courses](http://www.MachineLearning.courses)

Predictive Analytics World



Beyond life insurance, one top-five health insurance company predicts the likelihood that elderly insurance policy holders will pass away within 18 months, based on clinical markers in the insured's recent medical claims.

For more details, see the article, "Deathwatch: Five Reasons Organizations Predict When You Will Die": <http://www.predictiveanalyticsworld.com/patimes/deathwatch-five-reasons-organizations-predict-when-you-will-die/>

... and is there prediction after death? It turns out that death is not the final event to be predicted for a life. The Chicago Police Department predicts whether a murder can be solved. The department found that characteristics of a homicide and its victim help predict whether the crime will be solvable.

# Predictive Analytics for Insurance

- 1) Pricing and selection
- 2) Fraud detection
- 3) Marketing
- 4) Workforce analytics, e.g., agent scoring
- 5) Claims management
  - a) Fast-tracking claims
  - b) Channeling claims, e.g., by predicting severity

Eric Siegel @predictanalytic

[www.MachineLearning.courses](http://www.MachineLearning.courses)

Predictive Analytics World



Case Study from Esurance:

Predictive Claims Management: Identifying the Optimal Repair or Loss Channel - "We have significantly improved Esurance's claims cost by placing a decision engine in front of their process to identify the optimal repair or total loss channel. Patience and diligence were key in building credibility with Esurance, CCC senior management, and the data science team. In this presentation, we will discuss how we turned an idea into a model, a solution, and ultimately a team. Given the lessons learned, a second initiative was subsequently executed upon and brought to market in a quarter the time. Building vision - in addition to models - keeps the pipeline filled with development ideas for what is a growing team."

<http://www.predictiveanalyticsworld.com/sanfrancisco/2015/agenda.php#day2-415a>

# Machine Learning Leadership



A very particular end-to-end practice that forges a shrewd path to deployment.

Eric Siegel @predictanalytic

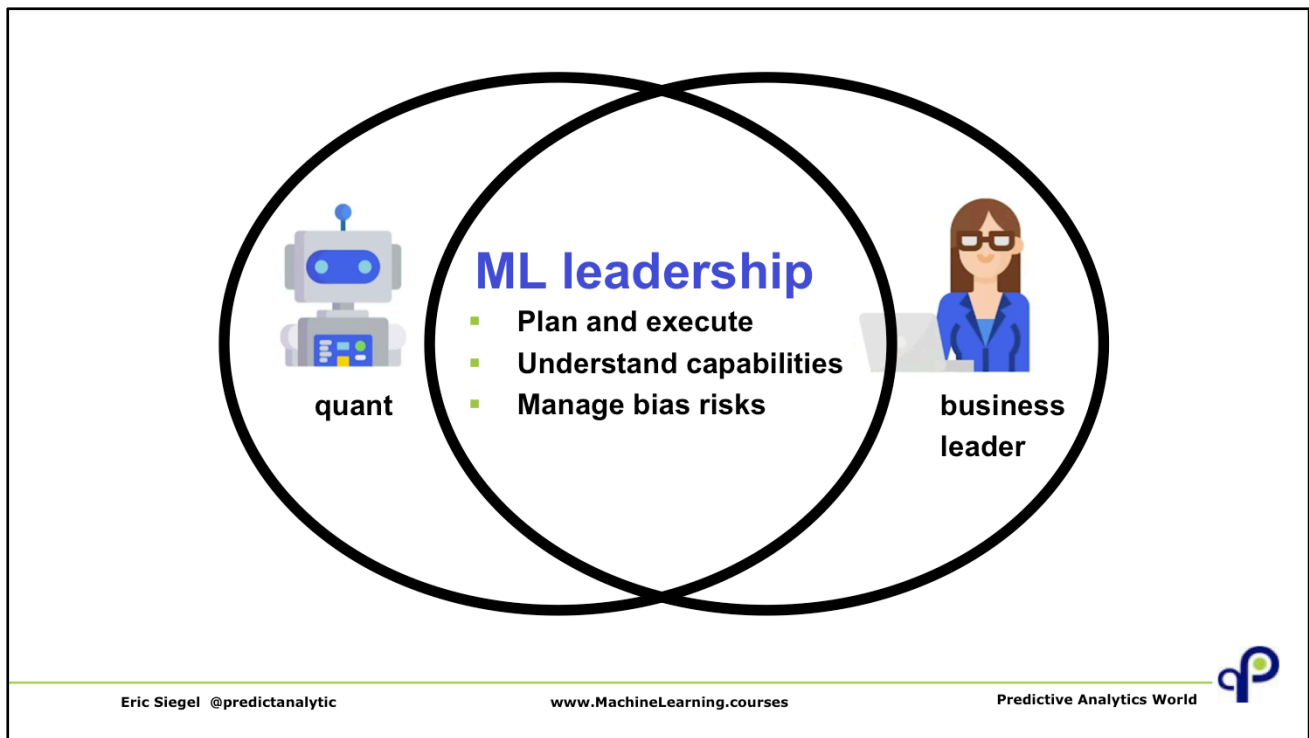
[www.MachineLearning.courses](http://www.MachineLearning.courses)

Predictive Analytics World



The right leadership practice is very often the one key missing ingredient for any machine learning project.





<https://www.predictiveanalyticsworld.com/machinelearningtimes/machine-learning-missing-link-business-leadership/12022/>

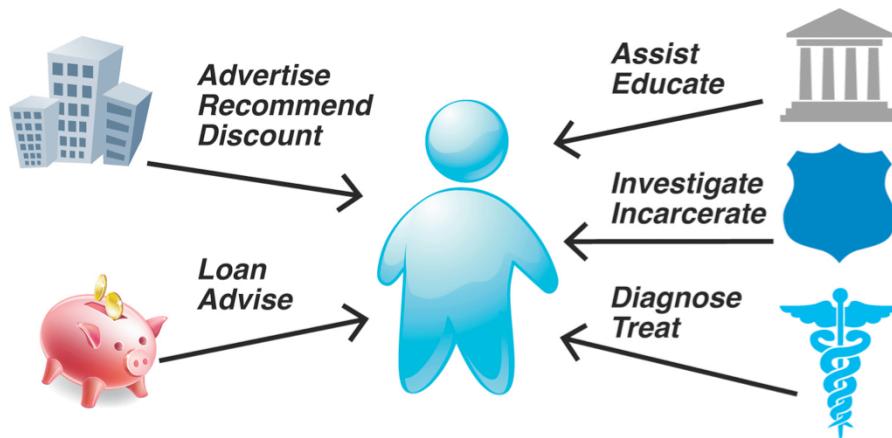
<https://hbr.org/2018/10/3-common-mistakes-that-can-derail-your-teams-predictive-analytics-efforts>

# Machine Learning Leadership

- 1) Business objective
- 2) Prediction objective
- 3) Data prep
- 4) Apply ML (TRAINING)
- 5) Deployment (SCORING)
- 6) Evaluate and maintain



Only by gaining buy-in on each one of these from not only data scientists but also business-side stakeholders – including executives and line-of-business managers – can a machine learning project stand a chance of being successfully deployed. Without leadership to facilitate the right collaboration, most predictive models succeed only quantitatively – the number-crunching side is good, but the project stalls at deployment. Great numbers of viable models never see the light of day.



Millions of decisions a day determine whom to ***call, mail, approve, test, diagnose, warn, investigate, incarcerate, set up on a date, and medicate.***



## ML Take-Aways:

- **Many applications**
  - Boosts operational effectiveness
  - Alleviates information overload
- **Learns from data**
- **Predict per *individual* – actionable**
- **The Data Effect:**

*Data is always predictive.*

- **The Prediction Effect:**

*A little prediction goes a long way.*





## Project leadership in-depth:

- Is machine learning strategic or tactical?
- How to pick your first project?
- How long does a project take?
- What staff do I need?
- How much data do I need?



ONLINE COURSE SERIES

## Machine Learning Leadership and Practice – *End-to-End Mastery*

Learn the state-of-the-art  
techniques and the business-  
side best practices.



[www.MachineLearning.courses](http://www.MachineLearning.courses)

### **Machine Learning Leadership and Practice – End-to-End Mastery**

This end-to-end, three-course will empower you to launch machine learning. Accessible to business-level learners and yet vital to techies as well, it covers both the state-of-the-art techniques and the business-side best practices.

<http://www.MachineLearning.courses>

# To Learn More: The Predictive Analytics Guide

[www.pawcon.com/guide](http://www.pawcon.com/guide)



Machine learning is the Information Age's latest evolutionary step.

The new electricity.

The most important general-purpose technology of our era.

**“The Information Age’s latest evolutionary step” – Eric Siegel**

Has emerged as a commonplace business practice necessary to sustain competitive advantage.

LinkedIn's first/second "Hottest Skills That Got People Hired" is "statistical analysis and data mining"