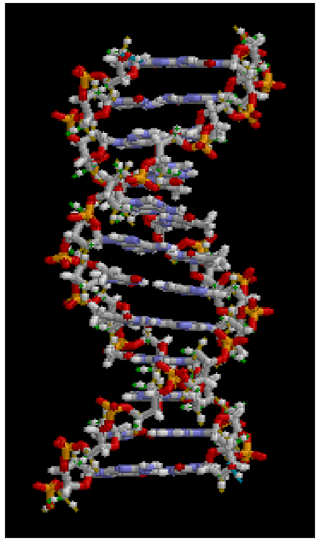
Artificial Intelligence and Machine Learning: What are they? How can we adapt to them?

Introduction to Machine Learning
St. Louis Actuaries Club 26-Jun-2018



Dave Snell, FALU, FLMI, ASA, MAAA, CLU, ChFC, ARA, ACS, MCP
technology evangelist, SnellActuarialConsulting - dave@ActuariesAndTechnology.com

MARYVILLE UNIVERSITY

Introductions

David L. Snell, FALU, FLMI, ASA, MAAA, CLU, ChFC, ACS, ARA, MCP
+1-314-278-8210 dave@ActuariesAndTechnology.com

PROFESSIONAL DEVELOPMENT

FALU, 2017; ASA, 1976; MAAA, 1986; FLMI, 1985; CLU, 1990; ChFC, 1994; ACS, 2001; ARA, 2003; MCP, 1998

EDUCATION

BSE, 1970 Major: Mechanical Engineering, Minors: Chemical, Civil, Electrical Engineering, Mathematics

WORK EXPERIENCE

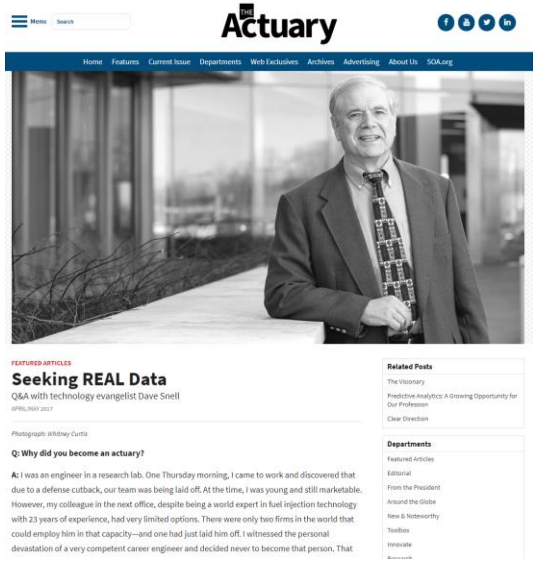
Technology Evangelist, Global Research and Data Analytics
RGA Reinsurance Company, Chesterfield, MO
(through Actuarial Innovations, LLC of St. Louis, MO since I am already retired from RGA) January, 2015 – May, 2017

Educate the RGA actuarial, underwriting, finance, and IT associates on the benefits and utilization of technology-based innovations and tools. Promote the use of complexity sciences - including machine learning, predictive analytics, genetic algorithms, behavioral economics, chaos theory, bioinformatics, sentiment analysis, recommender systems, etc.; and select or invent tools to facilitate the successful implementation of these concepts and tools for competitive business advantage. Nurture our young PhDs and other bright associates in the predictive analytics area; and mentor them in communication skills and career planning. Notable achievements include getting department associates published in several journals, and having them present at various actuarial, engineering, underwriting and other financial services conferences. I encouraged and nurtured one talented associate to attain his PhD in March of 2017 and am acknowledged in his thesis.

Technology Evangelist, Office of the Vice Chair
RGA Reinsurance Company, Chesterfield, MO
(through Actuarial Innovations) September, 2007 – December, 2014

Same job duties as above, but on a company-wide basis (until the retirement of the Vice Chair in December, 2014). During my time is this position I was co-inventor of, and wrote the bulk of our patent application for, a patent granted in 2014 (U.S. Patent 8775218) which combines machine intelligence with human intelligence for a synergistic outcome better than either could accomplish singly. Other patents have been granted in Japan, India, South Africa, and Canada.

VP, Asia-Pacific Technology
RGA Reinsurance Company, Chesterfield, MO July, 1998 – September, 2007
(Increasingly responsible actuarial and IT positions through the years, eventually leading and managing all technological development in Asia, Australia, and New Zealand. One notable achievement was the invention and development of an AI-based expert system, the Automated Underwriting Risk Assessment (AURA) system which has since been translated into a dozen languages (including Chinese) and has brought in over 100 Billion dollars of life reinsurance to RGA.


Menu Search
The Actuary
Home Features Current Issue Departments Web Exclusives Archives Advertising About Us SOA.org

FEATURED ARTICLES

Seeking REAL Data

Q&A with technology evangelist Dave Snell

APRIL/MAY 2017

Photograph: Whitney Curtis

Q: Why did you become an actuary?

A: I was an engineer in a research lab. One Thursday morning, I came to work and discovered that due to a defense cutback, our team was being laid off. At the time, I was young and still marketable. However, my colleague in the next office, despite being a world expert in fuel injection technology with 23 years of experience, had very limited options. There were only two firms in the world that could employ him in that capacity—and one had just laid him off. I witnessed the personal devastation of a very competent career engineer and decided never to become that person. That

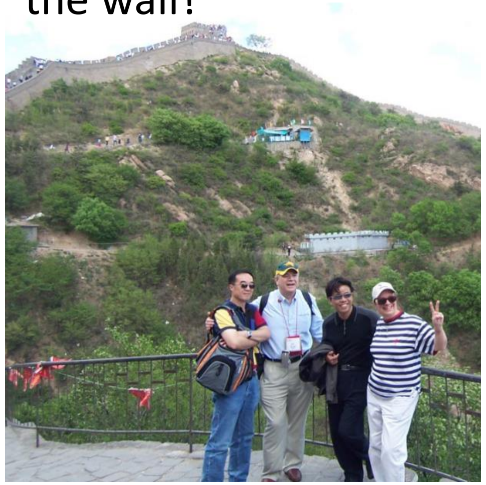
Related Posts
The Visionary
Predictive Analytics: A Growing Opportunity for Our Profession
Clear Direction

Departments
Featured Articles
Editorial
From the President
Around the Globe
New & Noteworthy
Toolbox
Innovate

Business casual in Beijing?

http://www.theactuarymagazine.org/seeking-real-data

Another brick in the wall?



A mob of fans

MARYVILLE UNIVERSITY

A Few Modest Goals for this first class

How is AI rapidly changing our world?
Why is Machine Learning suddenly feasible decades after the major algorithms were developed?
What implication does this have for insurance?

What are some of the major techniques involved (beyond typical regression and classification)?

tiny subset we can summarize today

- decision tree automation and random forests
- artificial neural networks, deep learning, and convolutional neural networks
- genetic algorithms

A few others we'll delve into later in the course (besides much going much deeper into these) are:
frequent itemsets, clustering, and other 'unsupervised' learning techniques, regression analysis,
Monte Carlo simulations, and behavioral economics, and programming languages for AI and ML.

How does machine learning differ from the classical tools and techniques that actuaries and IT professionals have been using?
What is the 'magic' behind these methods?

What are some case studies of successful past applications … and lost opportunities?
Where (and how) can we stay viable in the future?

What questions do you have?

Let’s talk
about
Data!

Machine Learning had to wait for ample and affordable data, computing power, and storage

According to an IDC study for EMC, humans have created 132 exabytes of data from the beginning of human civilization to year 2005.

2010: 1,200 exabytes … 2015: 7,900 exabytes … 2020: 40,700 exabytes of data (est.)

Source: Vernon Turner – Senior VP, Enterprise Infrastructure, Consumer, Network, Telecom and Sustainability Research, IDC – April, 2014
https://www.emc.com/leadership/digital-universe/2014iview/digital-universe-of-opportunities-vernon-turner.htm

Perspective:
1 byte – a single character
2 kilobytes – a typewritten page
5 megabytes – the complete works of Shakespeare
1 gigabyte – a pickup truck filled with paper filled with printed words
10 terabytes – the printed collection of the U.S. Library of Congress
2 petabytes – all U.S. academic research libraries
5 exabytes – all words ever spoken by human beings (printed)
http://highscalability.com/blog/2012/9/11/how-big-is-a-petabyte-exabyte-zettabyte-or-a-yottabyte.html

1 Billion hours of video viewed on Youtube every day
https://www.youtube.com/yt/about/press/ 27-Oct-2017

Machine Learning had to wait for ample and affordable data, computing power, and storage

As of 26-June-2018 you can buy more than 100,000 times the storage for under $250 and it will be about as large as your fingernail.

In 1957, the IBM 350 storage drive held 3.5 MB. You could rent it for $3,200 per month. The 350's cabinet is 5 feet long; 5 feet, 8 inches high and 2 feet, 5 inches wide.

Source: Wikipedia 24-Oct-2017
https://en.wikipedia.org/wiki/History_of_IBM_magnetic_disk_drives#IBM_350

Machine Learning had to wait for ample and affordable data, computing power, and storage

In 1979, I needed an additional 16 K Language Card for my Apple II (in order to program in UCSD Pascal). It cost me more than $300 for the 16 K of additional random access storage. There was no option then for a hard drive.

Today, you can buy a 4 Terabyte hard drive for under $100.

"Today, your cell phone has more computer power than all of NASA back in 1969, when it placed two astronauts on the moon." - Dr. Michio Kaku, 2014 (theoretical physicist, futurist, and popularizer of science; professor of theoretical physics at the City College of New York and CUNY Graduate Center.)

The wait is over!

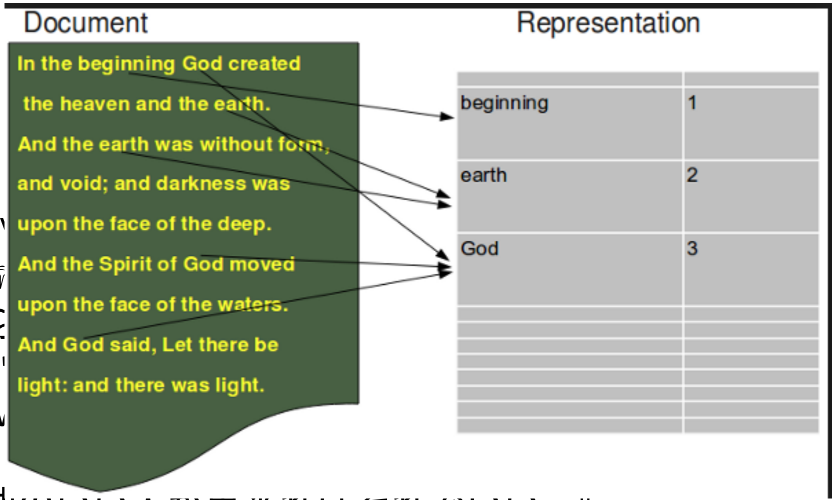
Cost	KB	Cost/KB	Cost/GB
$300 (1979)	16	$18.75	$18,750,000
$100 (2017)	4,000,000,000	$ 0.000000025	$ 0.025

Big Data is all around us – much publicly posted

1% sample of 332,900 tweets in 5 seconds

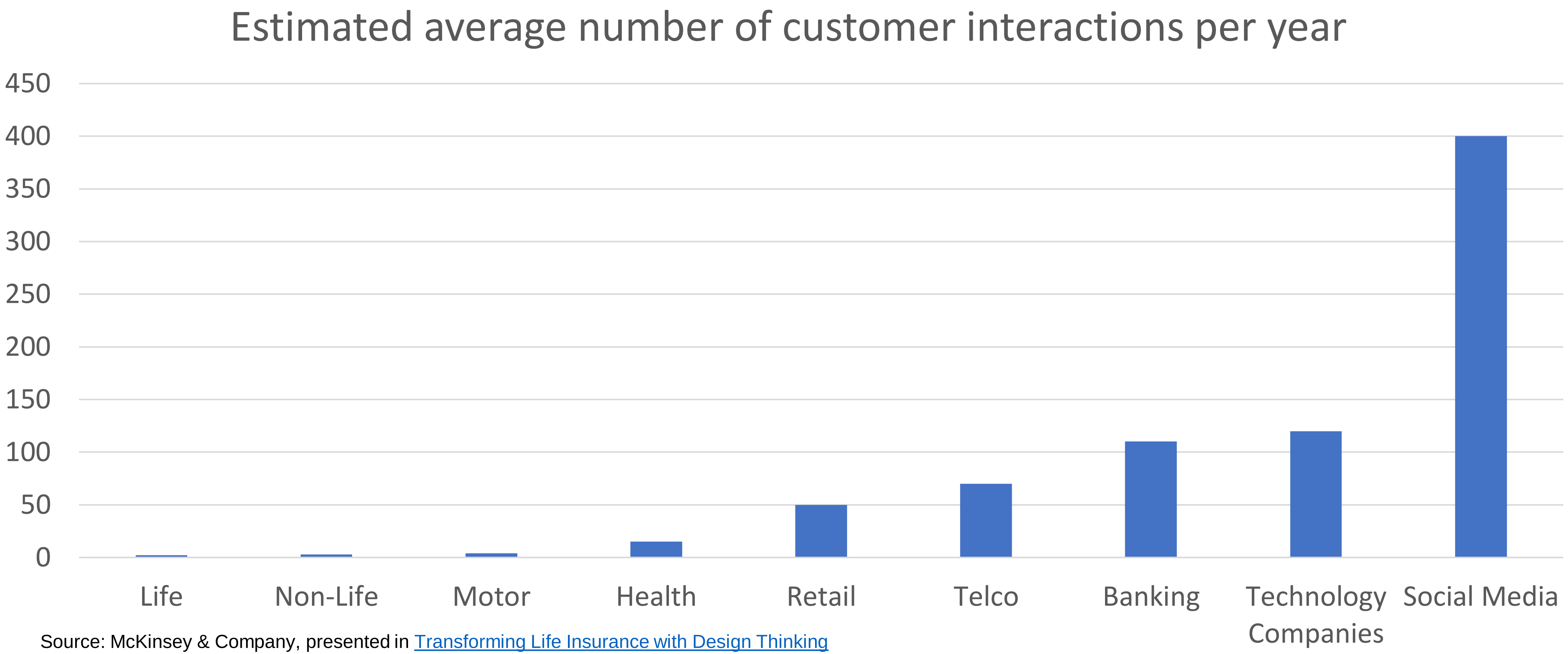
```
> proc.time()-ptm
   user  system elapsed
   0.08    0.00    5.02
>
> tweets.df <- parseTweets("tweets_sample.json")
```

- 332900 tweets have been parsed.
- > tail(tweets.df$text,20)
- [1] "RT @yuteesonyu: ไม่เห็นด้วยกับรูปนี้เลย ไม่ใช่คนไทยทุกคนที่คิดแบบนี้ แล้วก็ไม่ใช่ฝรั่งทุกคนที่คิดแบบนี้ คนไทยดีๆก็มี ฝรั่งแย่ๆก็มี https://…"
- [2] "Psychedelic Padded Pipe Pouch by https://t.co/GRpeEhB0n3 https://t.co/rDRSdbBN5v via @Etsy #hippy #weed #smoke #can
- [3] "RT @teed_chris: WISCONSIN,, TRUMPSTERS, AMERICANS, WE COME TOGETHER FOR A BATTLE TODAY, AND FOR OU
- [4] "@tabo_luv_ST 音だけ流れ続けて画面真っ暗～ｗｗｗ"
- [5] "@nozomieiei …知ってる"
- [6] "So much pain inside him.Immense betray from Yulin humans #StopYuLin4ever https://t.co/EZaxTDJ5q0"
- [7] "RT @sylvmic: Check out these awesome @5SOS headphones!! https://t.co/9hkaYaABwM #essential5SOS https://t.co/WfIzaxV
- [8] "RT @skywalkgrier: et le 3x01 qd il l'appel pr son anniv alors qu'il a perdu son humanité https://t.co/yNI7qIE0VU"
- [9] "猫をあやす棗さんが可愛すぎて歯磨き粉噴出した"
- [10] "こんな時間に腹減り"
- [11] "RT @tomozh: 大変だった時に使うハンコできた https://t.co/48VaQbVcpx"
- [12] "あっ"
- [13] "モイ！iPhoneからキャス配信中 - https://t.co/ccrG6sHn43"
- [14] "RT @KSeriesAD: พัคโบกอม ถ่ายแบบให้กับแบรนด์ MontBell คอลเลคชั่น S/S 2016 / หล่อ น่ารัก \xed��
- [15] "RT @SHXBL94_: ไม่ใช่คนที่โลกส่วนตัวสูงครับ ไม่ใช่คนที่เข้ากับคนยาก ตรงกันข้ามผมเข้ากับคนอื่นง่าย แต่ผมแค่เลือกคน
- [16] "RT @ARS_C_bot: 青「パクに土偶と埴輪の違いは解りますか？って聞いてみたら
で埴輪はこう(埴輪のポーズ)ですよね！』って答えられた。そういう話じゃない」"
- [17] "@kurooshiteru @tohruoikawa don't worry. Even in Japan I wouldn't have done that.
- [18] "Ladies https://t.co/ELNALcLYyu"
- [19] "【定期】すべての人に好かれる気はないし必要ないと思ってる。ごく少数の仲
- [20] "@june7845 고양이귀랑 꼬리랑 발 달고 고양이란제리랑 스타킹 입고 사진찍자"



Tweets captured by author May, 2016

Insurance companies have barely tapped into the amount of data available on our customers



Source: McKinsey & Company, presented in [Transforming Life Insurance with Design Thinking]

11-May-1997: Deep Blue Beats World Champion Chess Player

The 1997 rematch

Game #	White	Black	Result	Comment
1	**Kasparov**	Deep Blue	1–0	
2	**Deep Blue**	Kasparov	1–0	
3	Kasparov	Deep Blue	½–½	Draw by mutual agreement
4	Deep Blue	Kasparov	½–½	Draw by mutual agreement
5	Kasparov	Deep Blue	½–½	Draw by mutual agreement
6	**Deep Blue**	Kasparov	1–0	
***Result:* Deep Blue–Kasparov: 3½–2½**				

Source: Wikipedia as of 24-Oct-2017
https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov

MARYVILLE UNIVERSITY

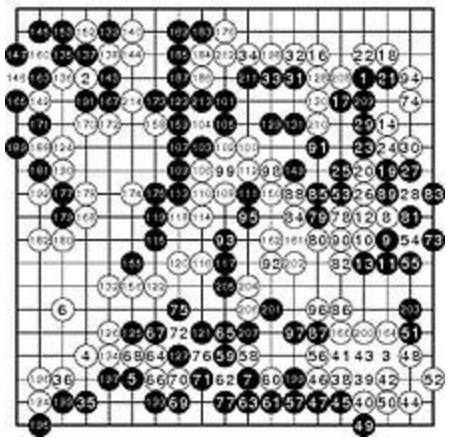
15-Feb-2011: Watson Beats Jeopardy Champions

Ken Jennings, Watson, and Brad Rutter in their *Jeopardy!* exhibition match.

Source Wikipedia as of 24-Oct-2017

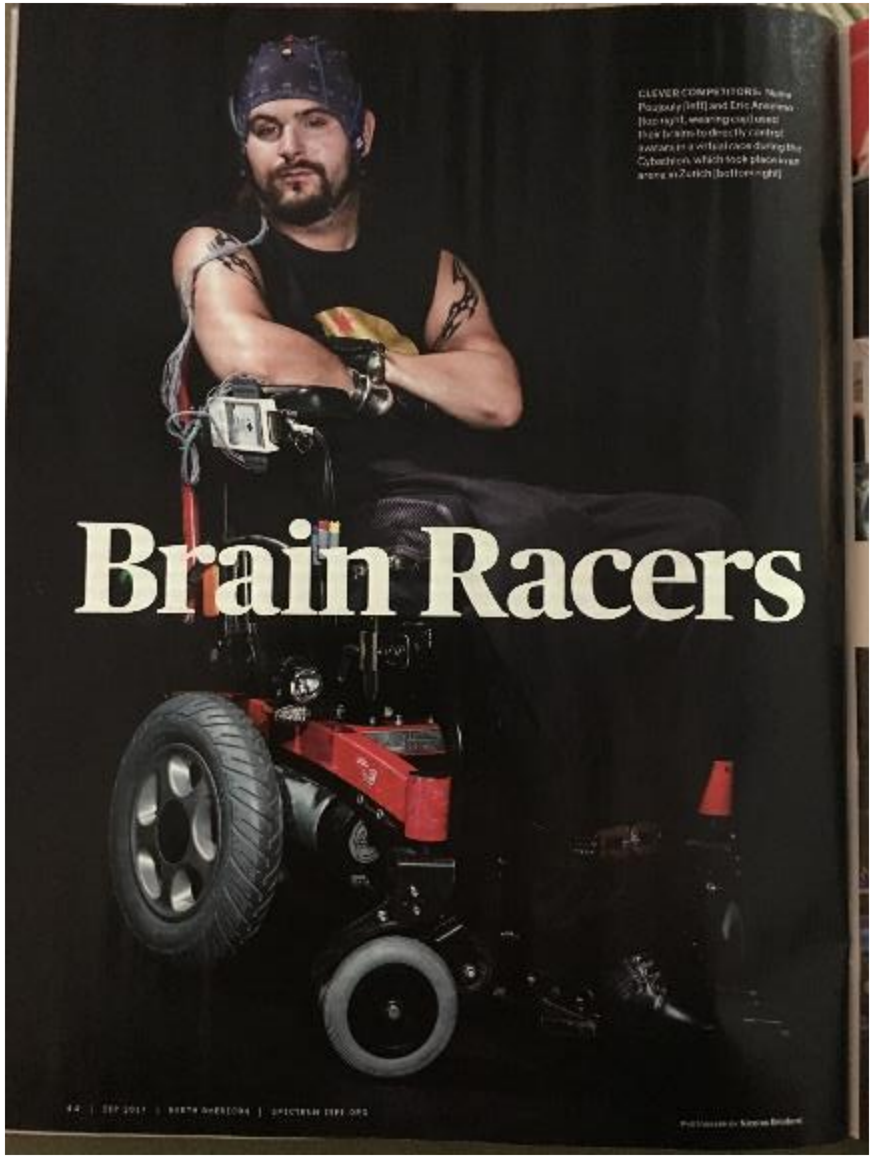
https://en.wikipedia.org/wiki/Watson_%28computer%29

15-Mar-2016: AlphaGo (Google DeepMind) Beats 18-time world champion Lee Sedol in Go



Game	Date	Black	White	Result	Moves
1	9 March 2016	Lee Sedol	**AlphaGo**	Lee Sedol resigned	186 Game 1
2	10 March 2016	**AlphaGo**	Lee Sedol	Lee Sedol resigned	211 Game 2
3	12 March 2016	Lee Sedol	**AlphaGo**	Lee Sedol resigned	176 Game 3
4	13 March 2016	AlphaGo	**Lee Sedol**	AlphaGo resigned	180 Game 4
5	15 March 2016	Lee Sedol[note 1]	**AlphaGo**	Lee Sedol resigned	280 Game 5

Source: Wikipedia 24-Oct-2017 https://en.wikipedia.org/wiki/AlphaGo_versus_Lee_Sedol



Sep-2017: AI helmets interpret brain waves and enable even quadriplegics to race in cyborg Olympics

IEEE Spectrum September, 2017

11-Sep-2017: Machine Learning was used at Georgia Institute of Technology to recreate game **code** (old Atari video games) **by watching the game** being played.

https://m.techxplore.com/news/2017-09-artificial-intelligenceuses-minutes-videogame-footage.html

18-Oct-2017: Google Deep Mind’s AlphaGo Zero Beats AlphaGo 100 games to zero

https://urldefense.proofpoint.com/v2/url?u=http-3A__www.bbc.com_news_technology-2D41668701&d=DwICAg&c=5uPv0lijNz76uSeaN5P0Zw&r=nkBVelrgsdY8RB9we12wjuanvgjtPHK9USWccjOv9NM&m=f_U2WRfiIOtw80O0kElpti1V9SR09KBHyTbVAujaXQ&s=mBZ7O9o9Szof_oRjuxxyPKvi8K5dK4BzzZx3Totkglo&e=

The AlphaGo program, devised by the tech giant's AI division, has already beaten two of the world's best players. It had started by learning from thousands of games played by humans.

But the new AlphaGo Zero began with a blank Go board and no data apart from the rules, and then played itself. Within 72 hours it was good enough to beat the original program by 100 games to zero.

AI rocks!

Machine Learning is amazing!

26-Oct-2017 **Sophia**, an intelligent humanoid robot, has been granted **citizenship** in Saudi Arabia.

- “I am very honored and proud for this unique distinction,” the robot said onstage. “This is historical to be the first robot in the world to be recognized with a citizenship.”
- CNBC's Andrew Ross Sorkin interviews Sophia, a humanoid robot, about the future of artificial intelligence at a Future Investment Institute panel in Saudi Arabia. Link to YouTube interview with Sophia:

  https://www.youtube.com/watch?v=S5t6K9iwcdw

Watch from 2:08 to 5:29 minutes

You may also enjoy her interview with Jimmy Fallon from 25-Apr-2017: https://www.youtube.com/watch?v=Bg_tJvCA8zw

7-Dec-2017 In Just 4 Hours, Google's AI Mastered All The Chess Knowledge in History

- "After being programmed with only the rules of chess (no strategies), in just four hours AlphaZero had mastered the game to the extent it was able to best the highest-rated chess-playing program Stockfish."
- "In a series of 100 games against Stockfish, AlphaZero won 25 games while playing as white (with first mover advantage), and picked up three games playing as black. The rest of the contests were draws, with Stockfish recording no wins and AlphaZero no losses."
- Source: https://futurism.com/4-hours-googles-ai-mastered-chess-knowledge-history/


YOU KEEP USING THAT WORD
But What is AI?
But What is Machine Learning?
slide idea from Jim Guszcza, US chief data scientist, Deloitte Consulting
I DO NOT THINK IT MEANS WHAT YOU THINK IT MEANS

What is Artificial Intelligence?
What is Machine Learning?

They are not magic!

"Any sufficiently advanced technology is indistinguishable from magic."
– Sir Arthur C. Clarke

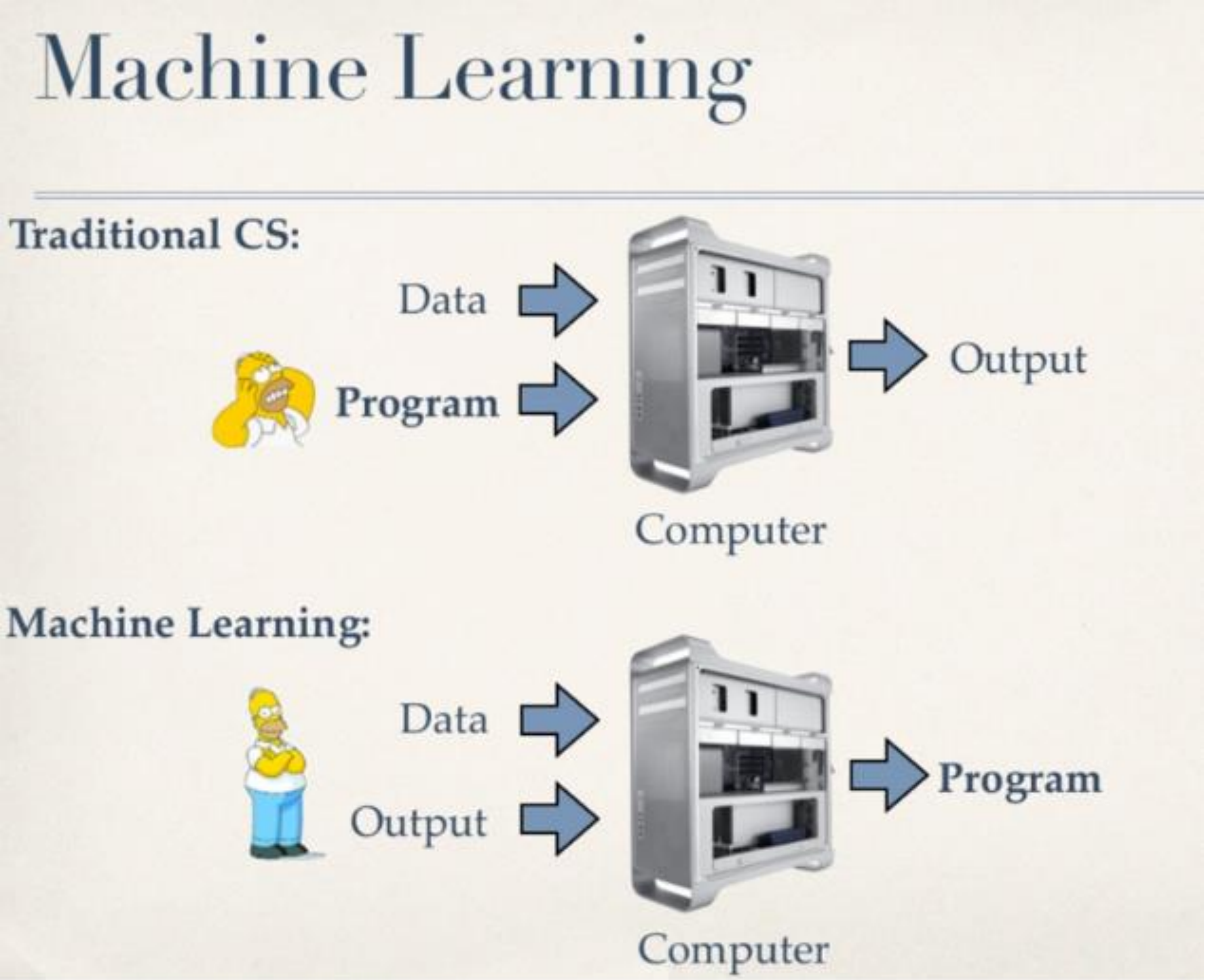
Forbes: 20-Jan-2016

Data Scientist tops list of ***best jobs*** in United States

Original picture by author at FIRST Robotics Competition, St. Louis, MO – April 2014

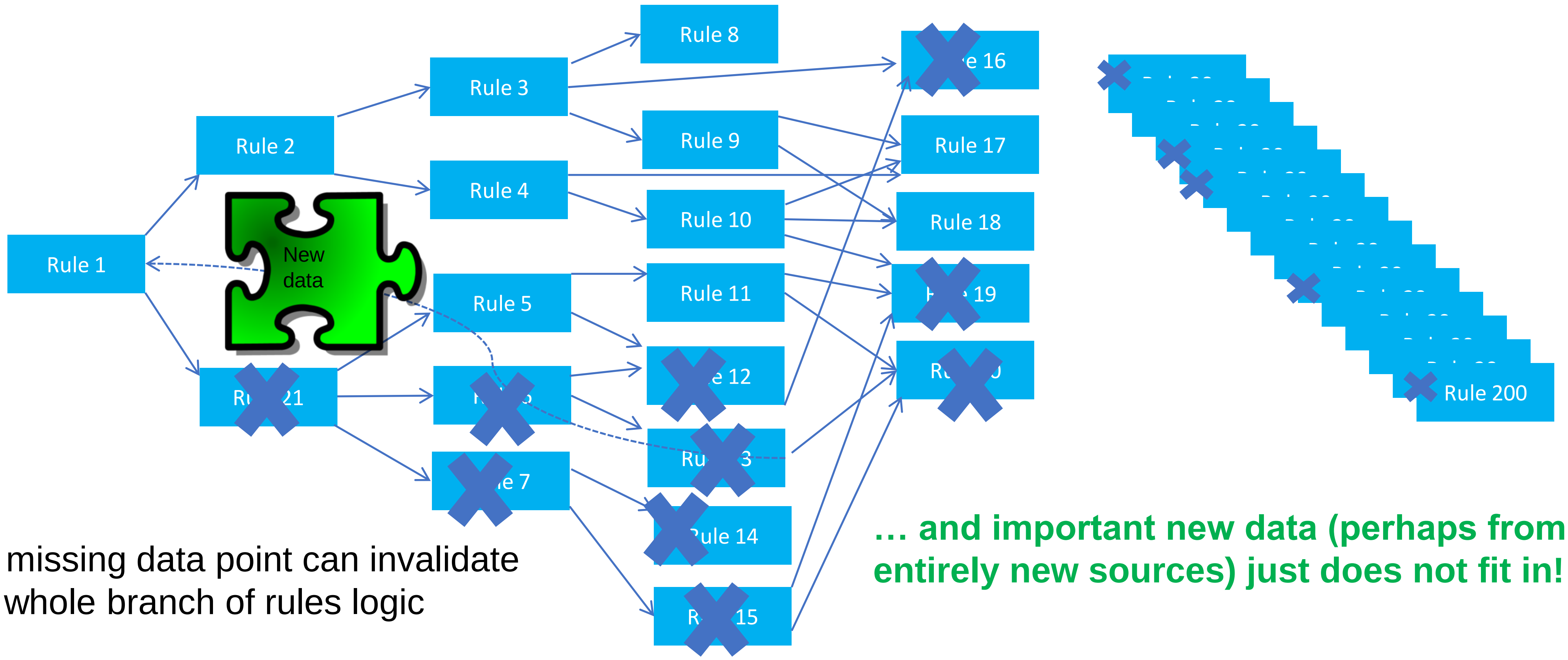
Overview of Machine Learning

Machine Learning



Source: Machine Learning in practice – common pitfalls, and debugging tricks, by Kilian Weinberger, Associate Professor, University of Washington

Traditional rule trees lack flexibility



A missing data point can invalidate a whole branch of rules logic

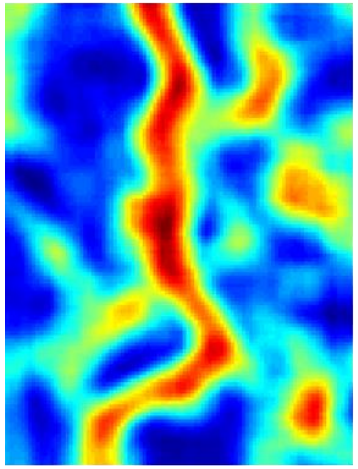
… and important new data (perhaps from entirely new sources) just does not fit in!

Images by Dave Snell

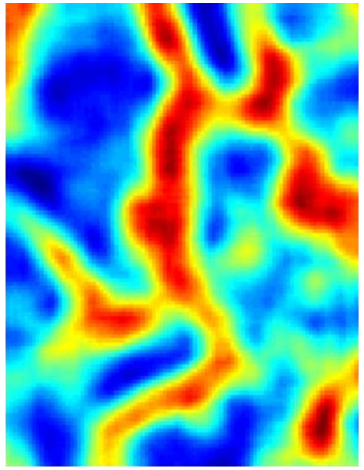
Pattern Matching is more flexible

Each scored case becomes a potential match

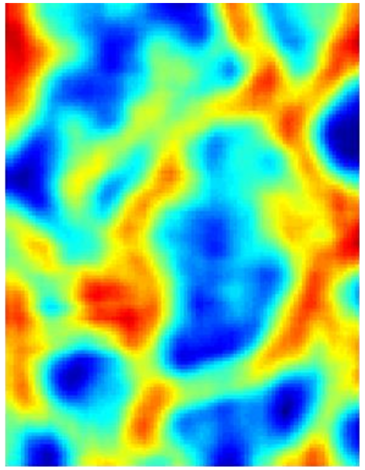
The fit may be perfect, or approximate, with a measure of closeness



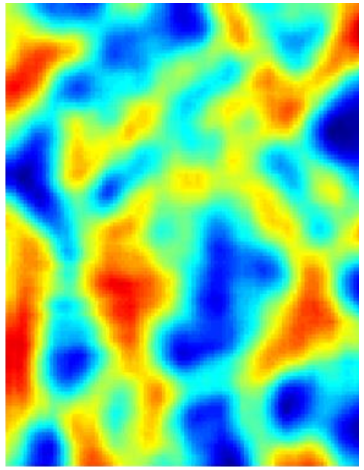
1



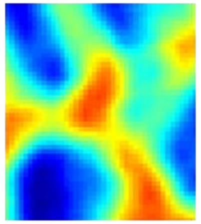
2



999



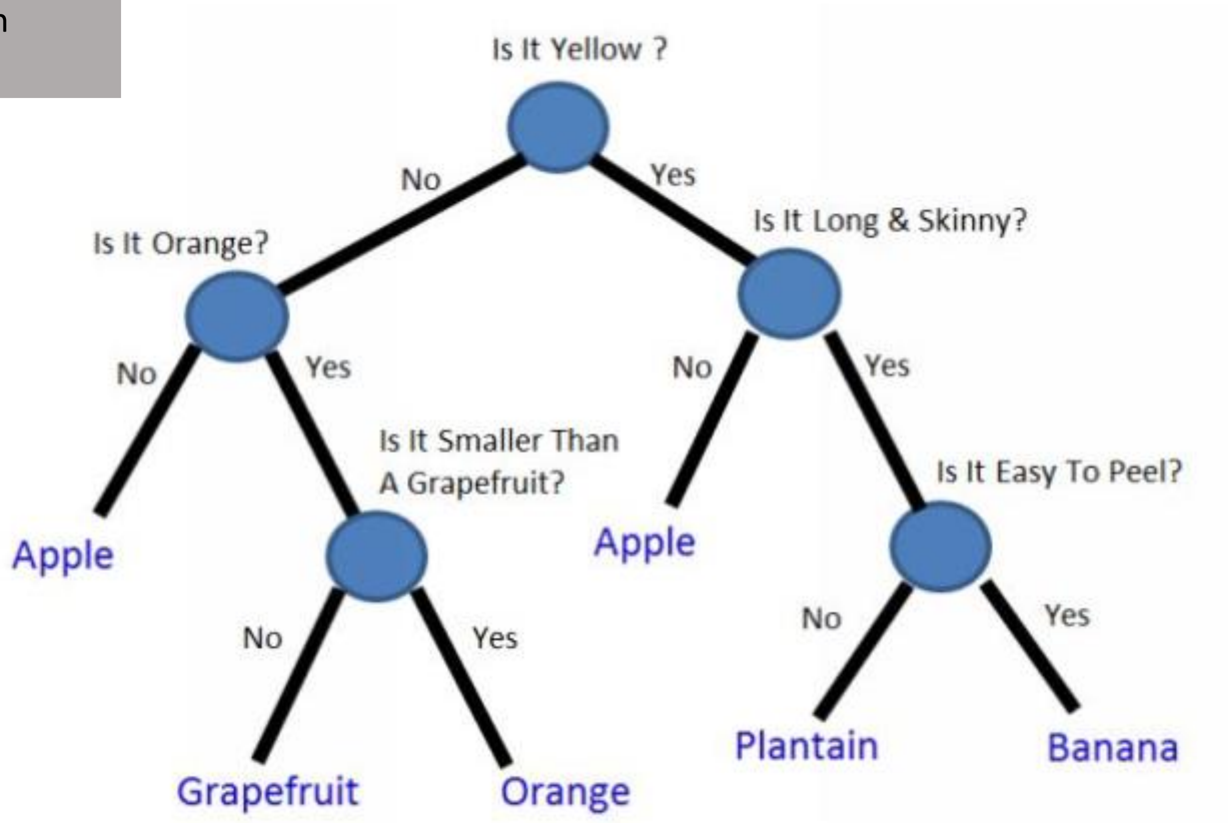
1000



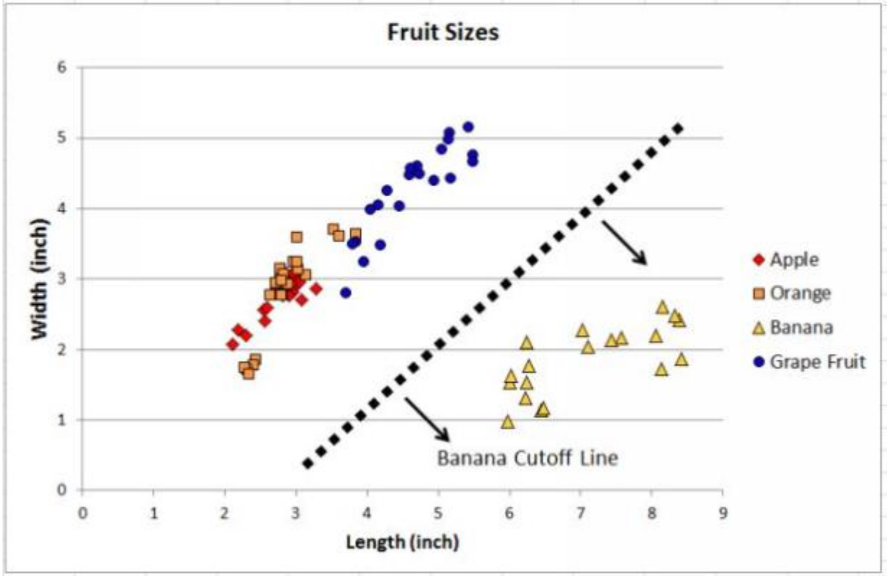
Case input pattern

images by Dave Snell

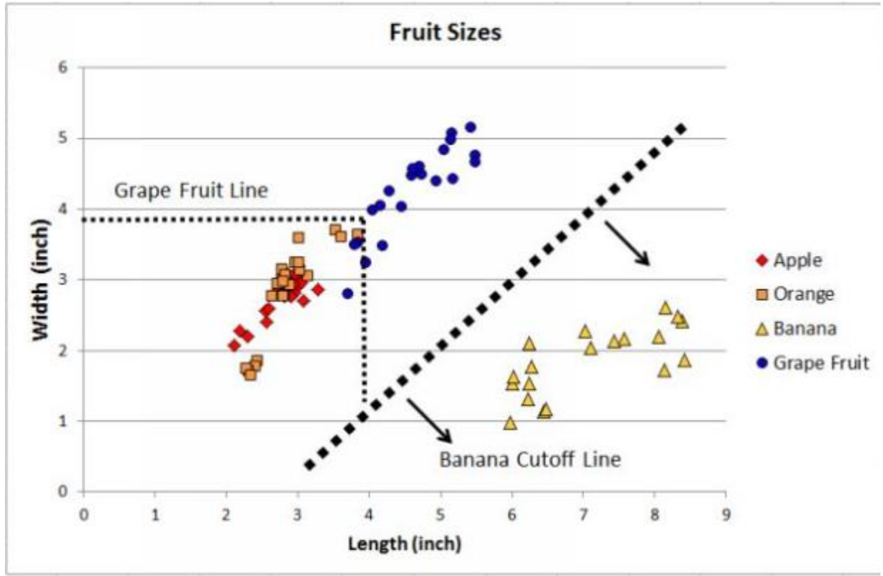
Decision tree example from *Machine Learning with Random Forests and Decision Trees* by Scott Hartshorn (available on Amazon for $2.99)



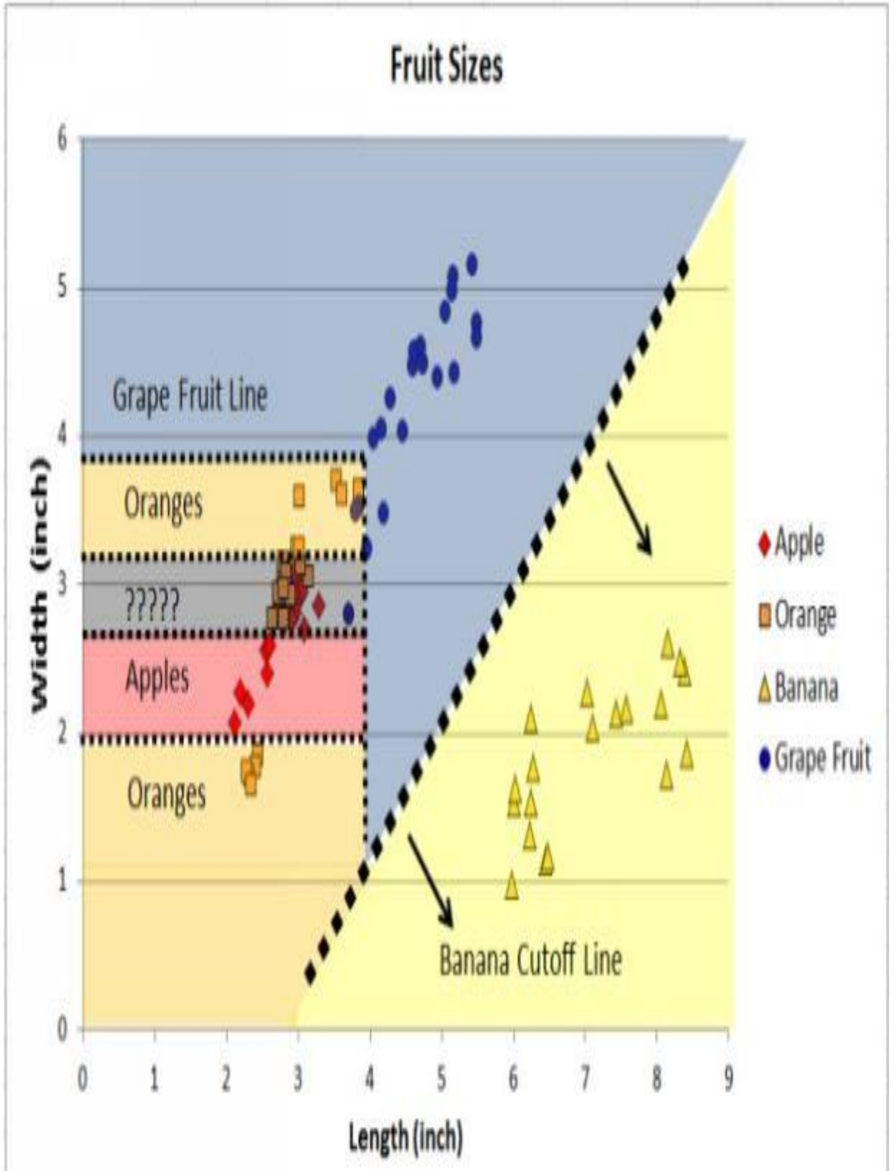
Simple decision tree to classify whether a fruit is an Apple, Grapefruit, Orange, Plantain, or Banana
Note: this example assumes those are the only possibilities. Thus, a lime would be classified as an Apple.



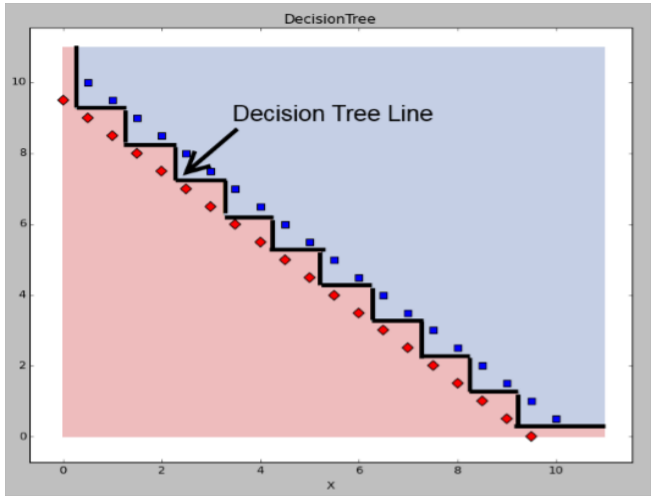
This is how a human might create a set of rules to segregate and identify the four types of fruit.

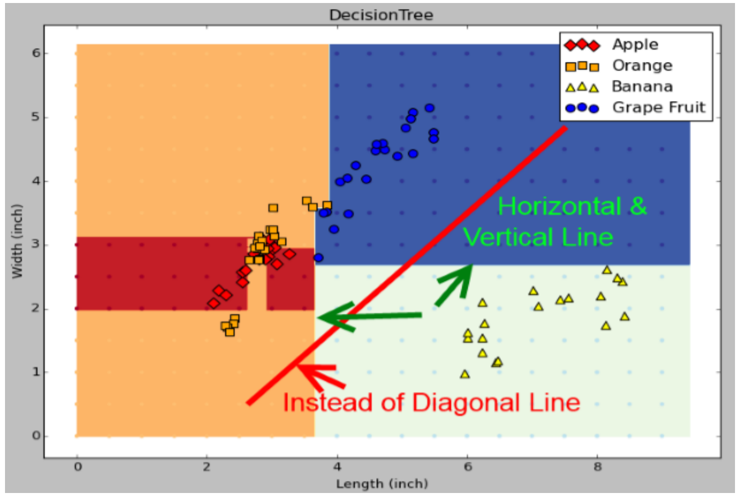


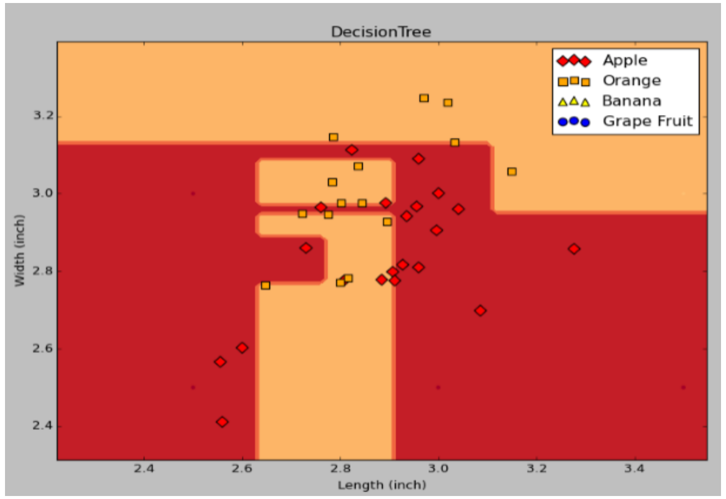
Quantitatively, the splits are placed where the Gini impurity (or other measure of fit) is minimized.



Decision tree example from *Machine Learning with Random Forests and Decision Trees* by Scott Hartshorn

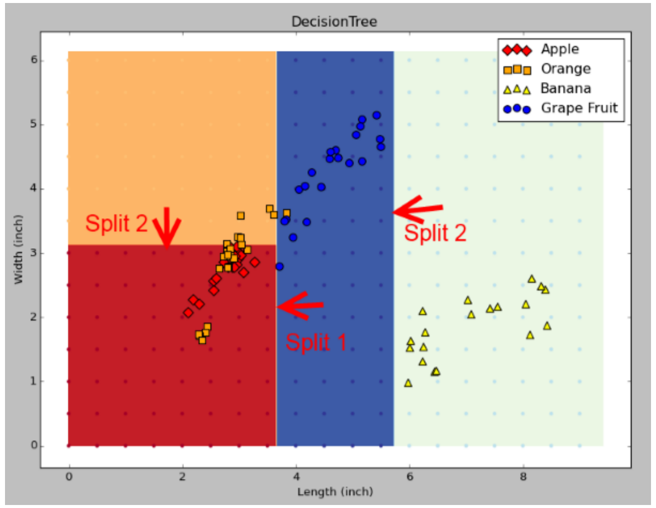




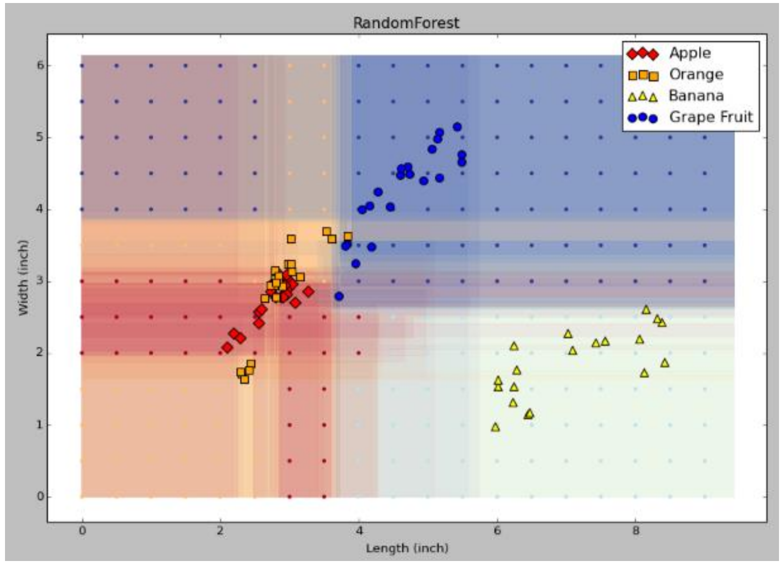


Unable to split diagonally, the decision tree algorithm must simulate the split with several vertical and horizontal splits.

In its attempt to classify ALL of the training data, an individual tree will be prone to overfitting. It is difficult to distinguish between true data, and noise.



A random forest is a collection of many trees, where *each tree has been given only a subset of the data and the parameters*. The impact of noise is greatly attenuated.



Decision tree example from *Machine Learning with Random Forests and Decision Trees* by Scott Hartshorn

How Many Trees Should You Have In Your Forest?

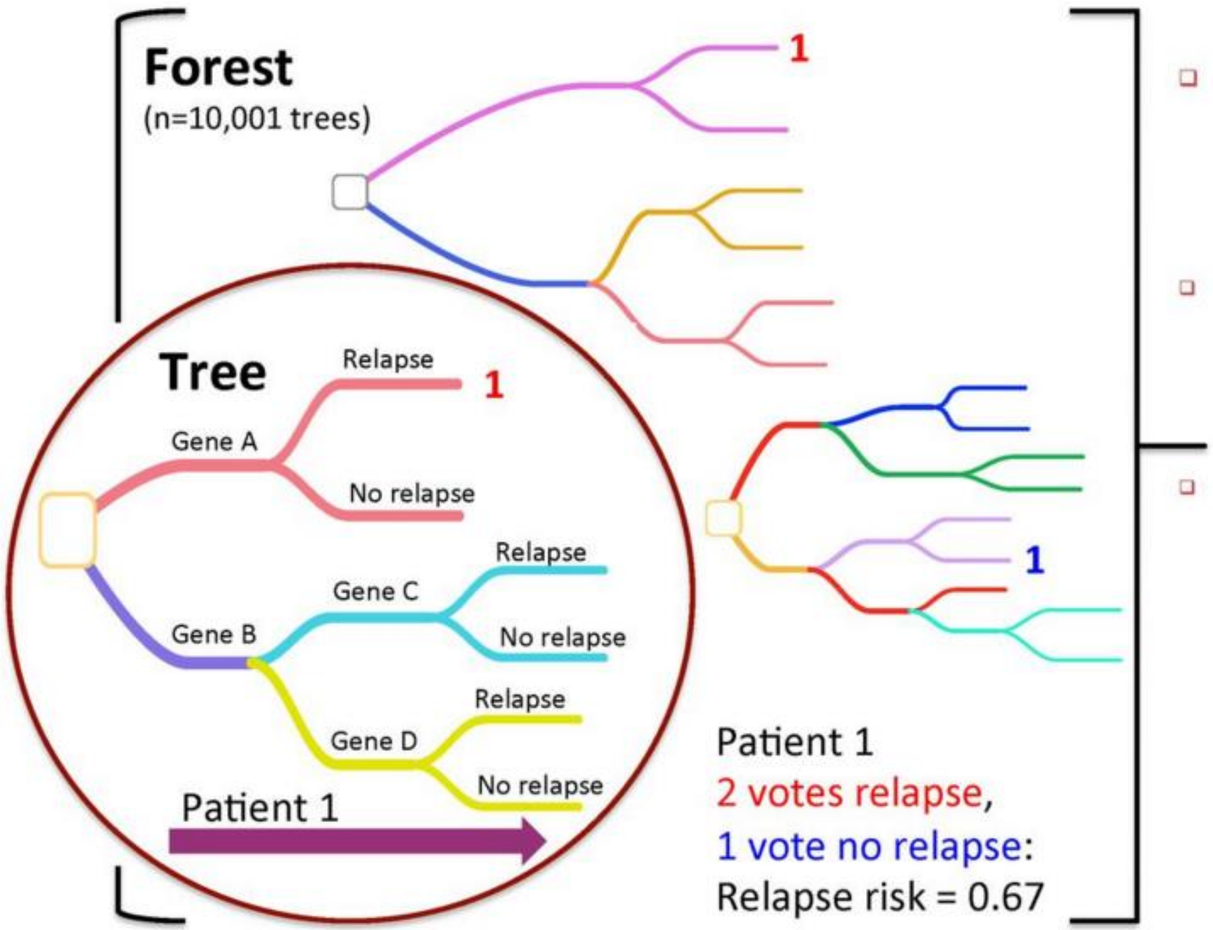
Advice from *Machine Learning With Random Forests And Decision Trees: A Visual Guide For Beginners*, by Scott Hartshorn

More trees are usually better because they will do more to smooth out abnormalities in the data. But that is true only up to a point. This is an area of diminishing returns, where each additional tree will have less benefit than the one before it. Eventually the benefit will plateau, and more trees will not help very much.

The decision on how many trees to have in the Random Forest becomes a tradeoff dependent on your problem and your computing resources. Going from 10 trees to 50 trees might improve your results significantly, and not add very much time. Going from 1,000 trees to 50,000 trees might add substantial time without improving your results very much.

Using 100 trees in the Random Forest is often a good place to start. Increase later (after tuning other parameters)

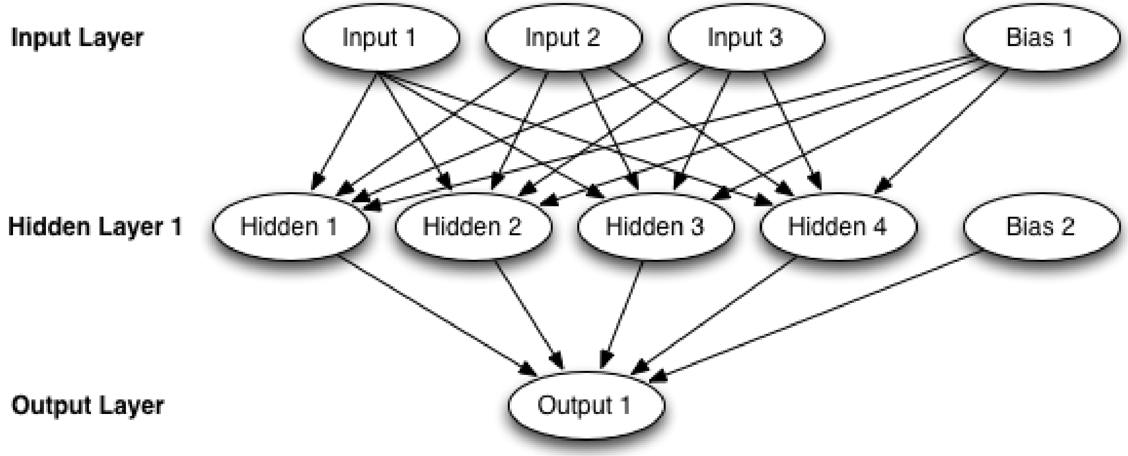
Overview of Random Forests

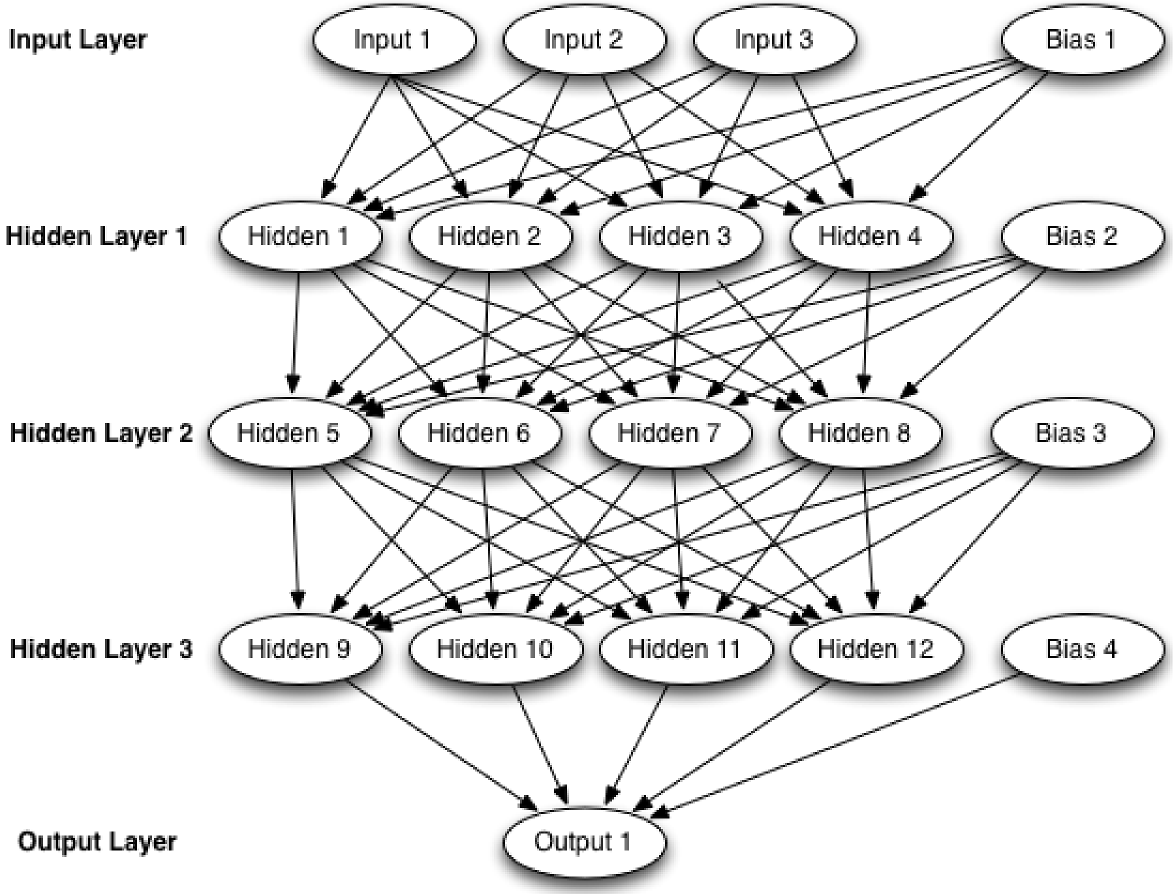


- **Creates a forest of many binary decision trees**
- Each Patient traverses each tree until it reaches a terminal node
- At the terminal node each tree casts a vote (eg. "relapse"); the proportion of relapse votes from all votes is that patient's predicted relapse risk

Images from Wikipedia and designated in public domain

Artificial Neural Networks





Images used with permission from: Heaton, J. (2015). *Artificial Intelligence for Humans, Volume 3: Deep Learning and Neural Networks*. St. Louis, MO: Heaton Research, Inc, 1505714346.

Very good, but still not perfect

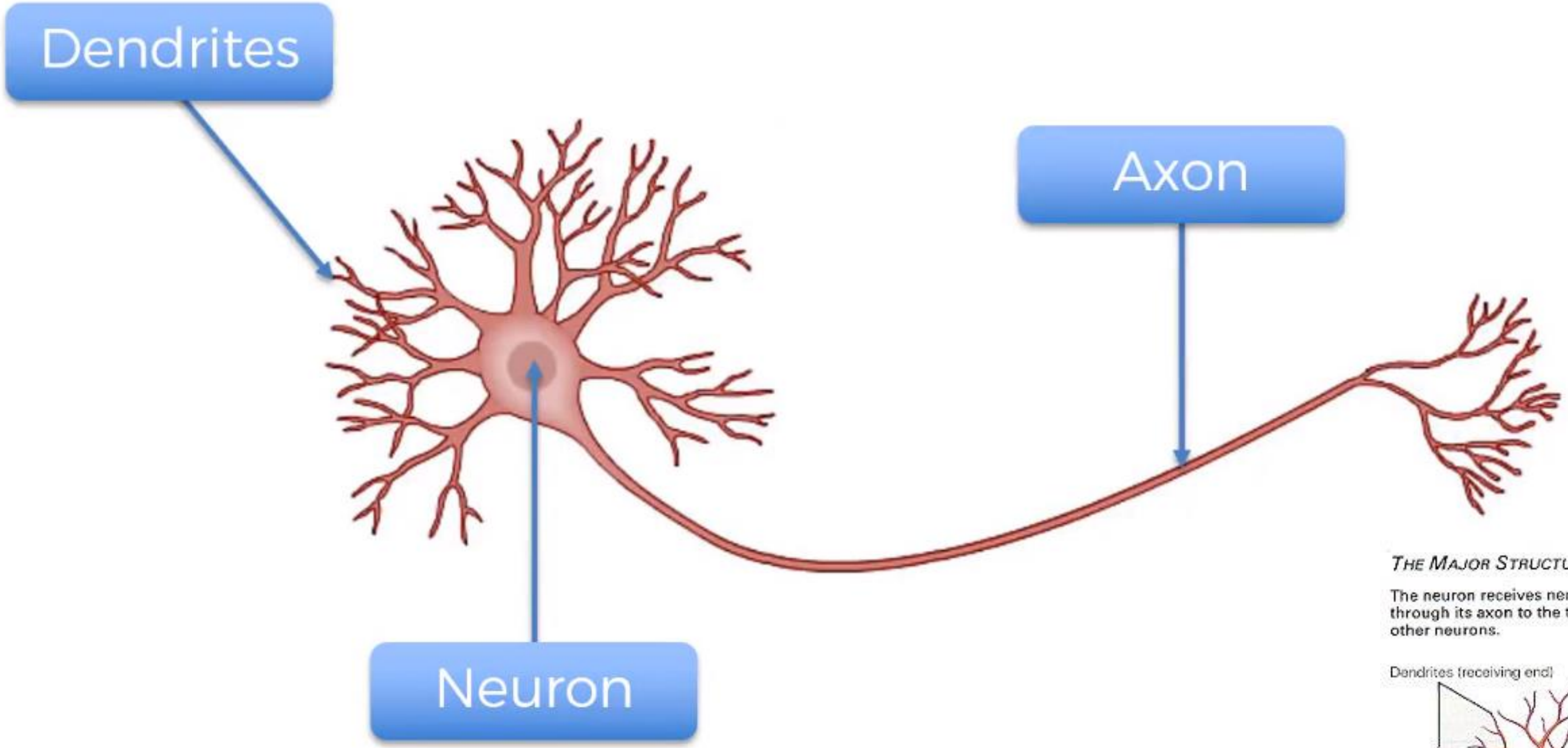
Convolutional Neural Networks provide the front end (for Deep Neural Networks) to enable facial recognition and autonomous vehicles.

MARYVILLE UNIVERSITY

By MICHELLE CASTILLO / CBS NEWS / *February 12, 2013, 3:02 PM*

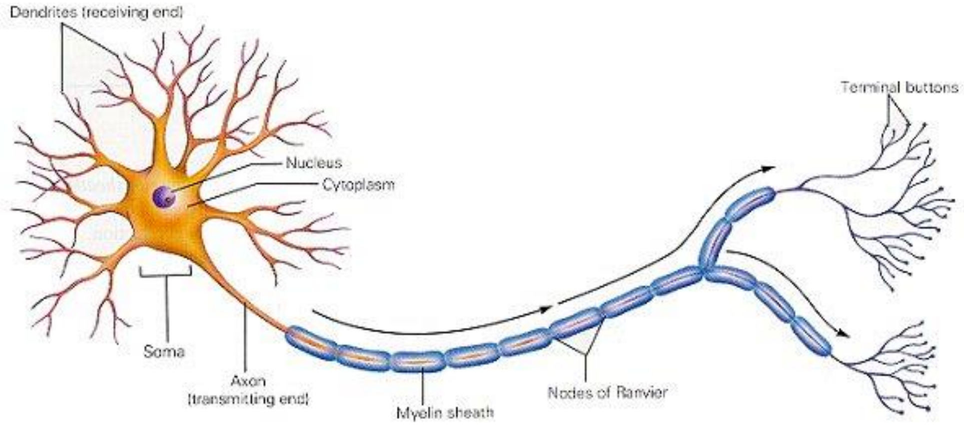
Can you spot the gorilla in this CT scan? Most radiologists couldn't

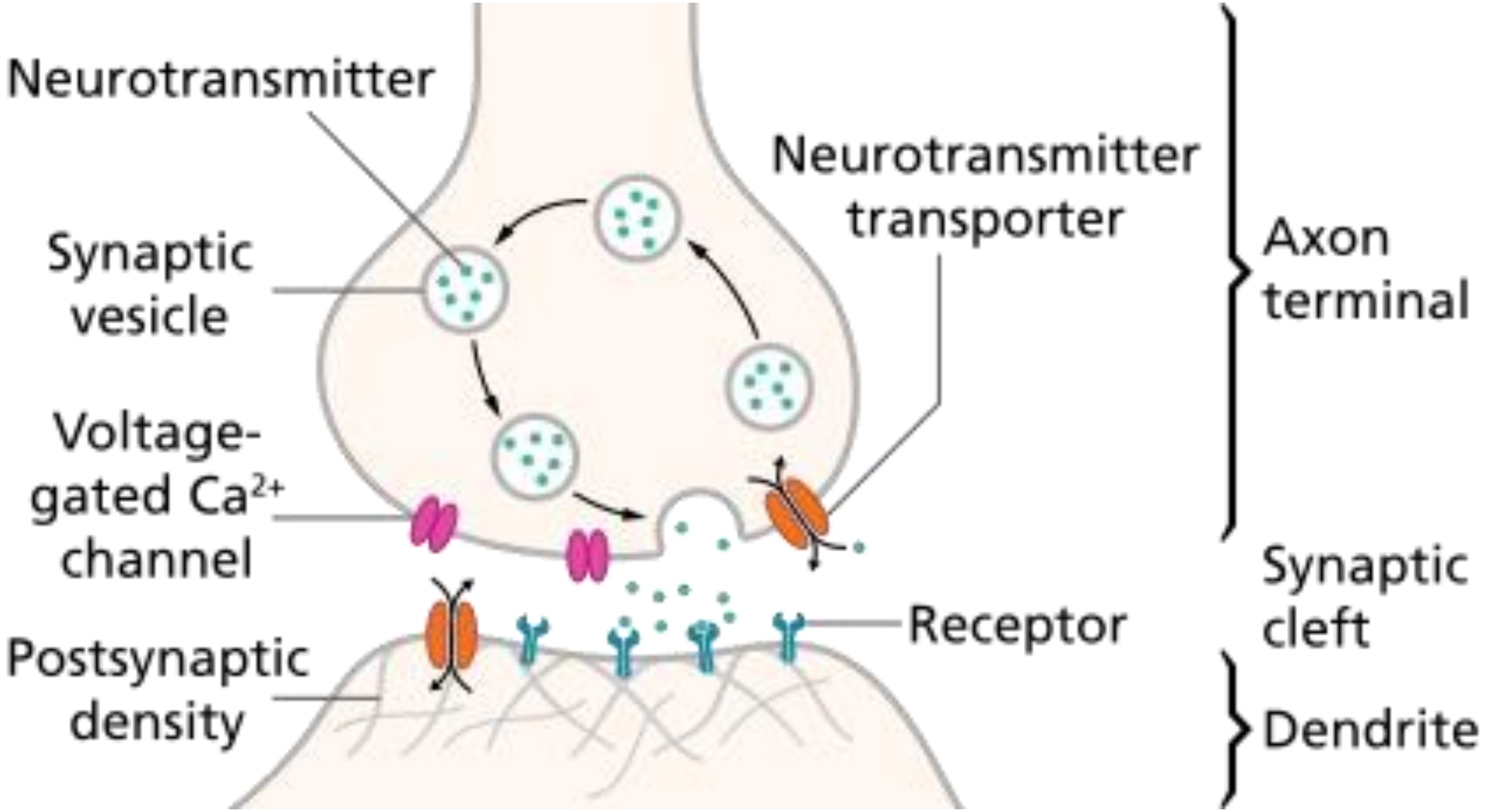
Do you see the gorilla in this scan? / **MELISSA VO, AND JEREMY WOLFE/PSYCHOLOGICAL SCIENCE,TRAFTON DREW**

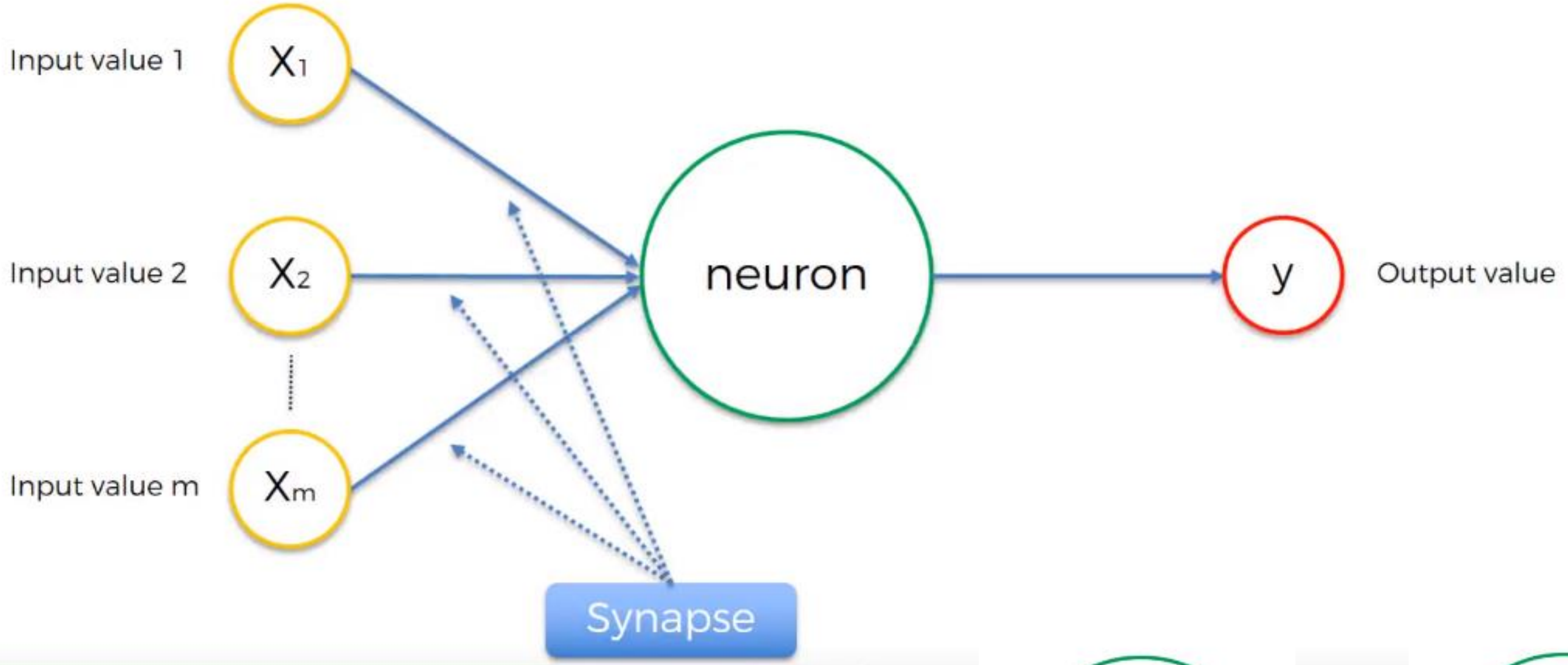


The Major Structures of the Neuron

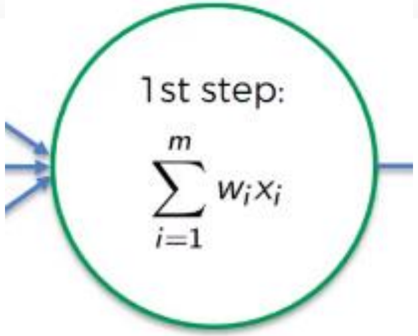
The neuron receives nerve impulses through its dendrites. It then sends the nerve impulses through its axon to the terminal buttons where neurotransmitters are released to stimulate other neurons.



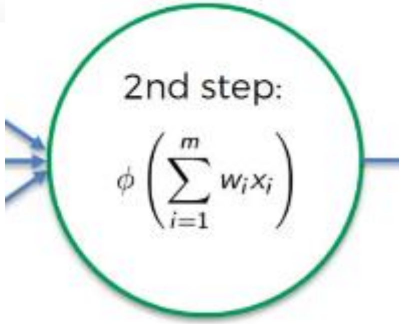

Neurotransmitter
Synaptic vesicle
Voltage-gated Ca2+ channel
Postsynaptic density
Neurotransmitter transporter
Receptor
Axon terminal
Synaptic cleft
Dendrite



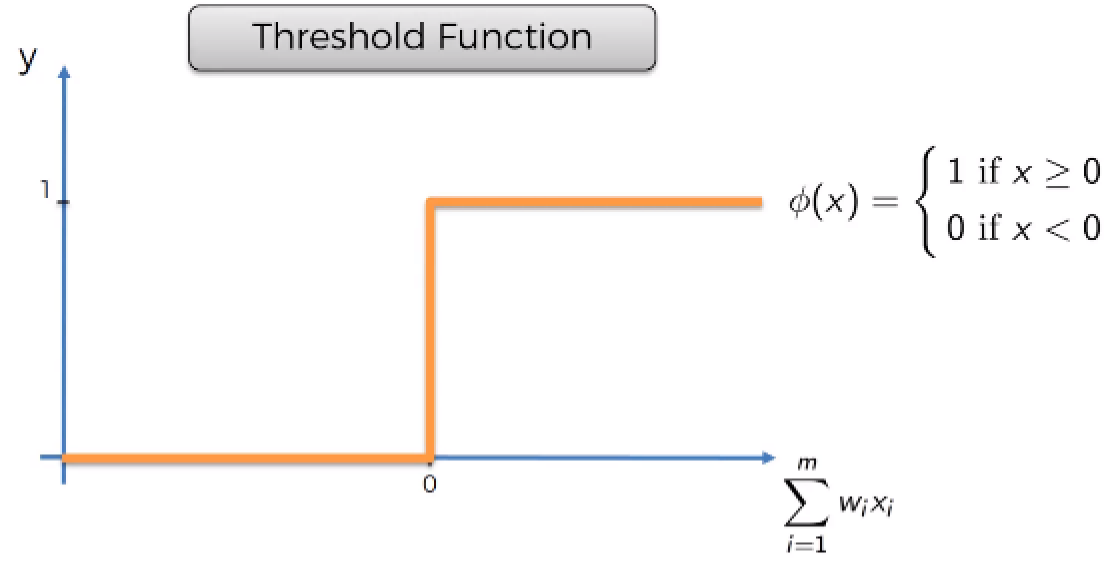
Input value 1
X_1
Input value 2
X_2
Input value m
X_m
neuron
y
Output value
Synapse

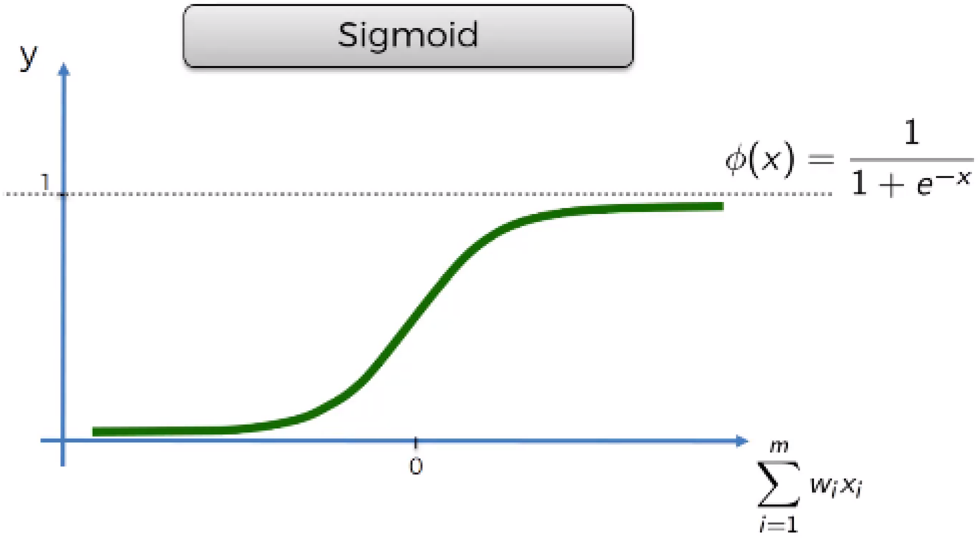


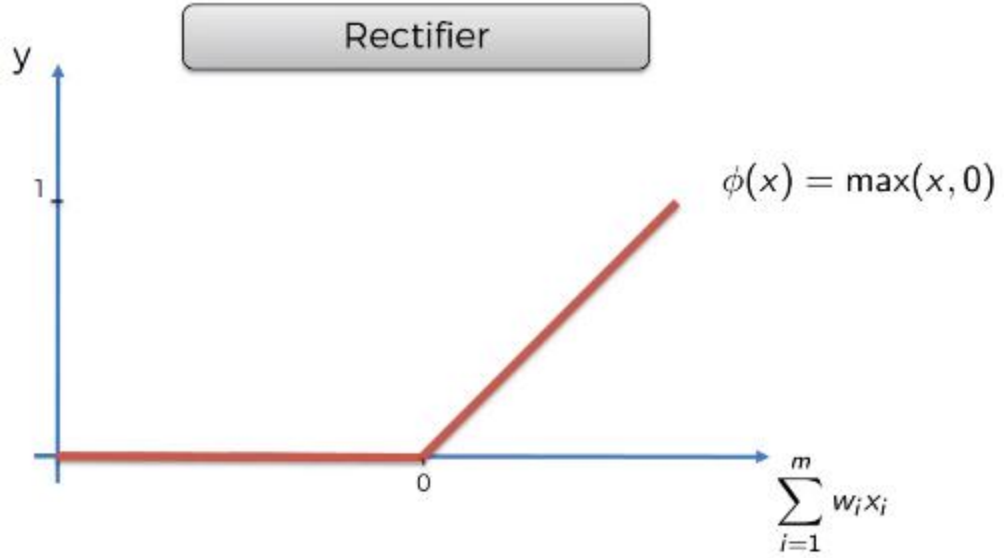
1st step:
$\sum_{i=1}^{m} w_i x_i$

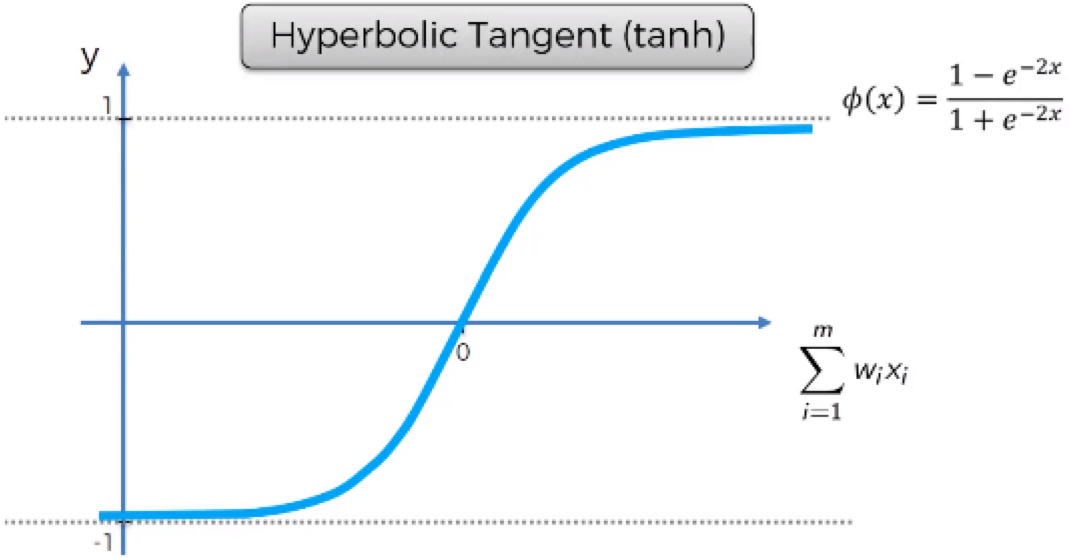


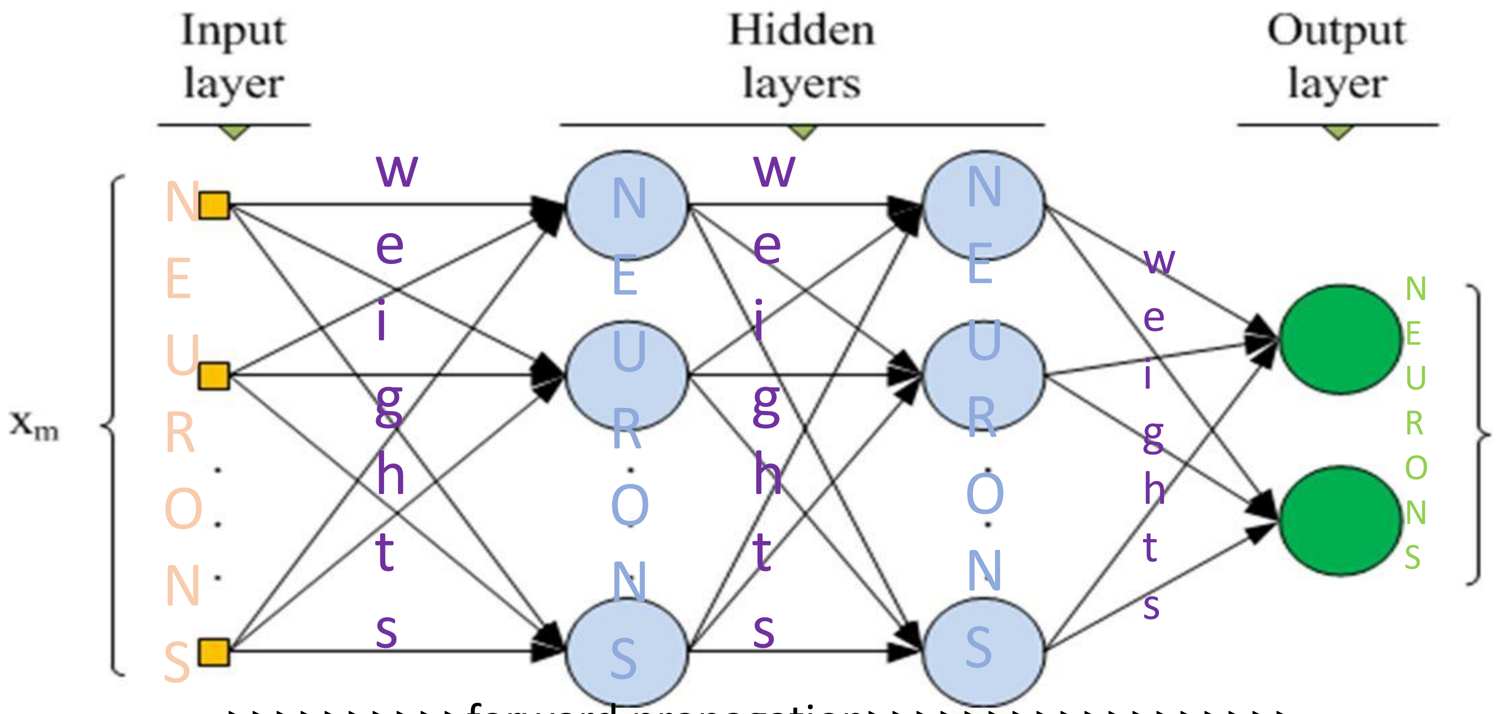
2nd step:
$\phi\left(\sum_{i=1}^{m} w_i x_i\right)$


Threshold Function
y
1
0
$\phi(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{cases}$
$\sum_{i=1}^{m} w_i x_i$


Sigmoid
y
1
0
$\phi(x) = \frac{1}{1 + e^{-x}}$
$\sum_{i=1}^{m} w_i x_i$


Rectifier
y
1
0
$\phi(x) = \max(x, 0)$
$\sum_{i=1}^{m} w_i x_i$


Hyperbolic Tangent (tanh)
y
1
0
-1
$\phi(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$
$\sum_{i=1}^{m} w_i x_i$



$\hat{Y}_m$ (output values)

Y_m = actual values

>>>>>>>>>>>forward propagation>>>>>>>>>>>>>>>>>>

<<<<<back propagation<<<<<<<<<<<<<<<<<<<<

Cost Function, $C = ½ \text{Sum}[(Y – \hat{Y})^2]$

Randomly initialize weights to small numbers (but not zero)

Forward propagate vector of first observation to get predicted result $\hat{Y}$. Each feature goes into one node of input layer.

Compute and minimize cost function, then update weights according to how much they are responsible for errors (and multiplying by the learning rate). This is called back propagation.

Continue to update weights after each observation (reinforcement learning) or after a batch of observations (batch learning)

An epoch occurs when all of the training set has passed through the process. Redo for as many epochs as desired.

Overview of Predictive Analytics Techniques

Deep Learning

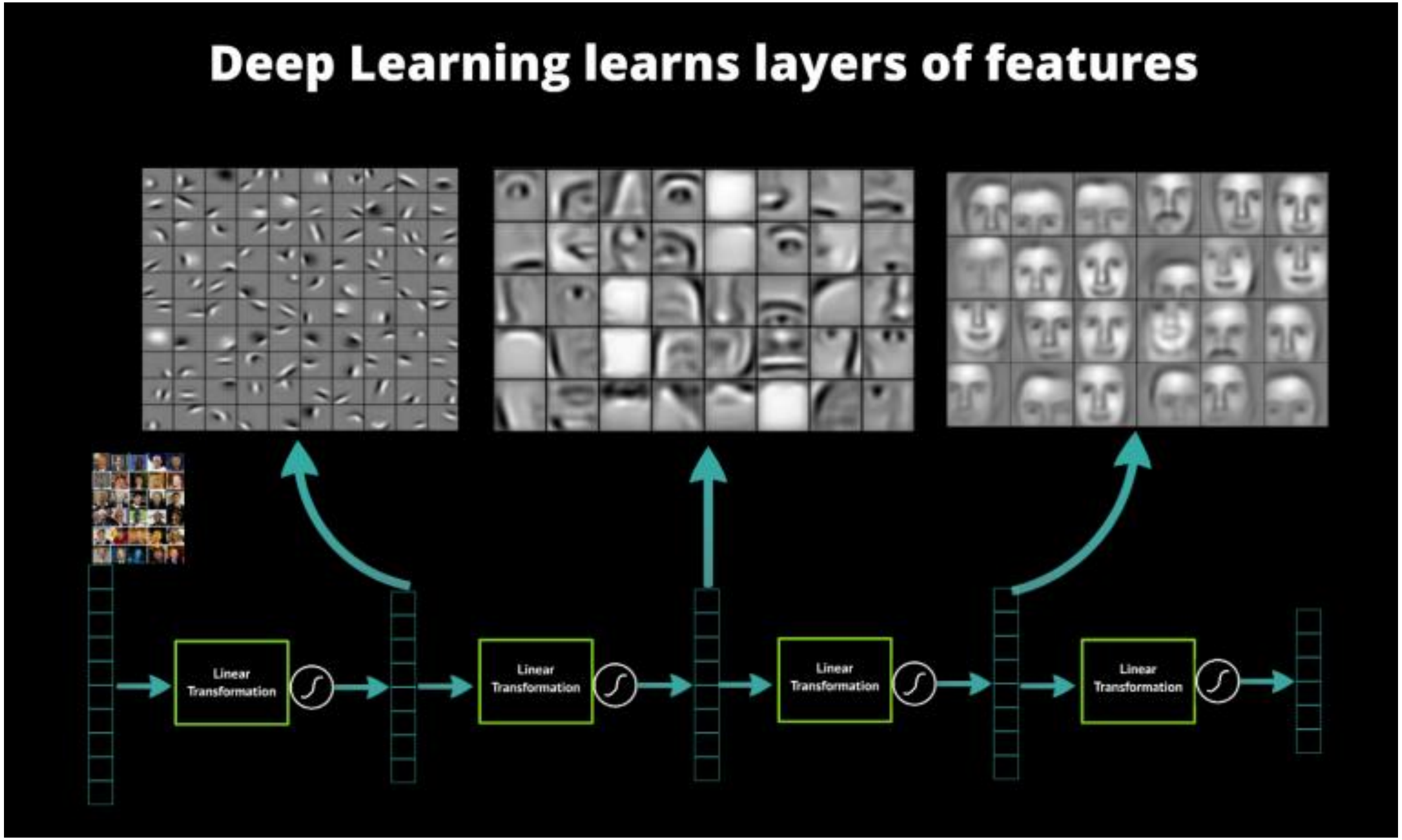
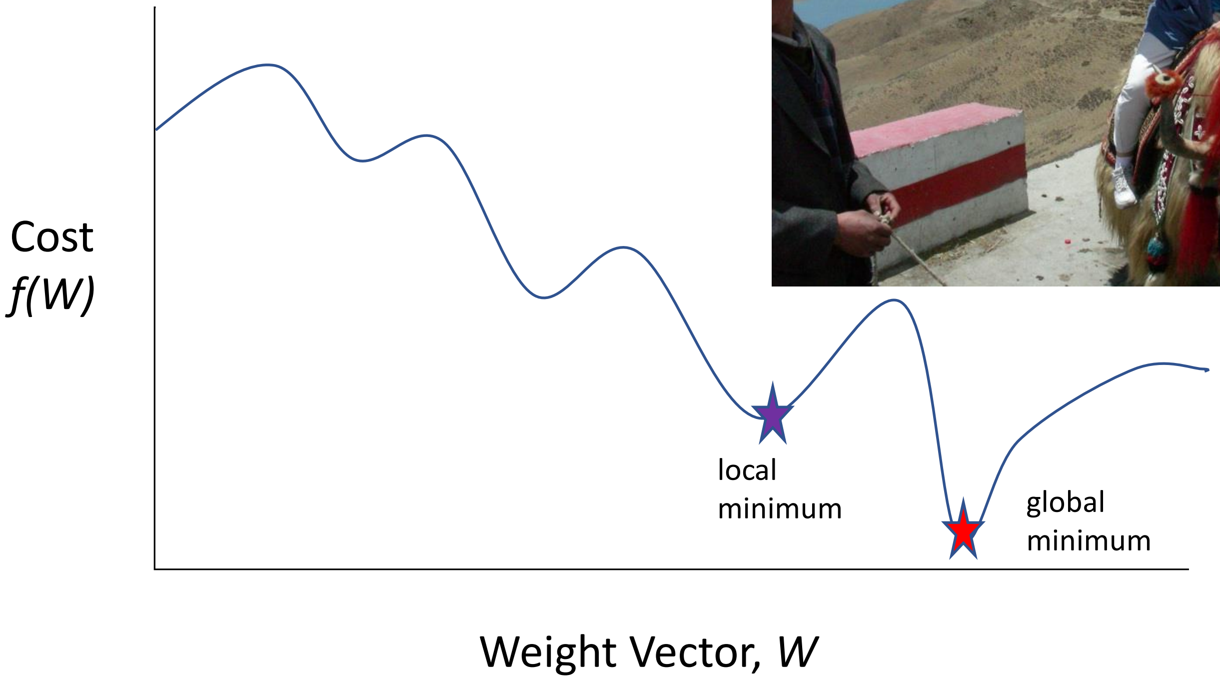


Image used with permission from DataRobot®: https://www.datarobot.com/blog/a-primer-on-deep-learning/

Simple Gradient Descent is great when you have a convex cost function; but what if you have a situation with several local minima? That's when Stochastic Gradient Descent is used to test several bands within the range of the cost function, and get to better solutions without the n
millions … or trillions) of trial and err



A free book to explain Deep Learning in more detail:
http://neuralnetworksanddeeplearning.com/

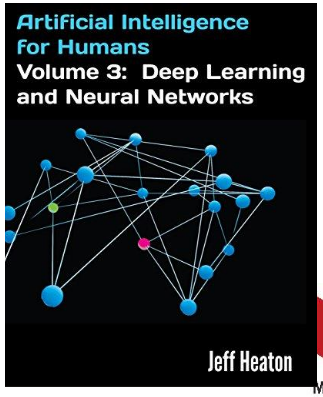
An interactive tool to help you understand the neural network process:

//www.heatonresearch.com/aifh/vol3/xor_online.html

re training a neural network for an XOR operator. You
ee literally every value that is calculated in the neural
ork to get each weight.

ey Heaton FLMI, ARA
Lead Data Scientist

A gentle introduction to Deep Learning is Jeff's book:



A best seller on Amazon

No current programming language is ideal for data science (DS), and as DS increases in popularity, there will be replacements that differ considerably from R and Python.

Source for charts: Kdnuggets 30-Aug-2017
https://www.kdnuggets.com/2017/08/python-overtakes-r-leader-analytics-data-science.html

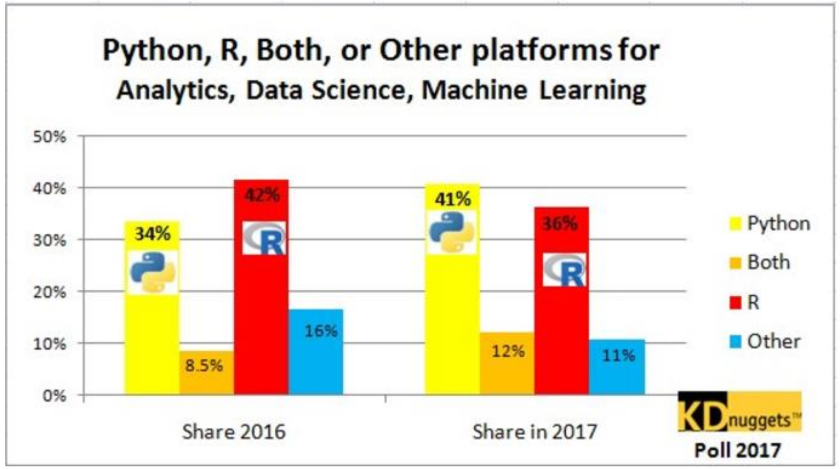
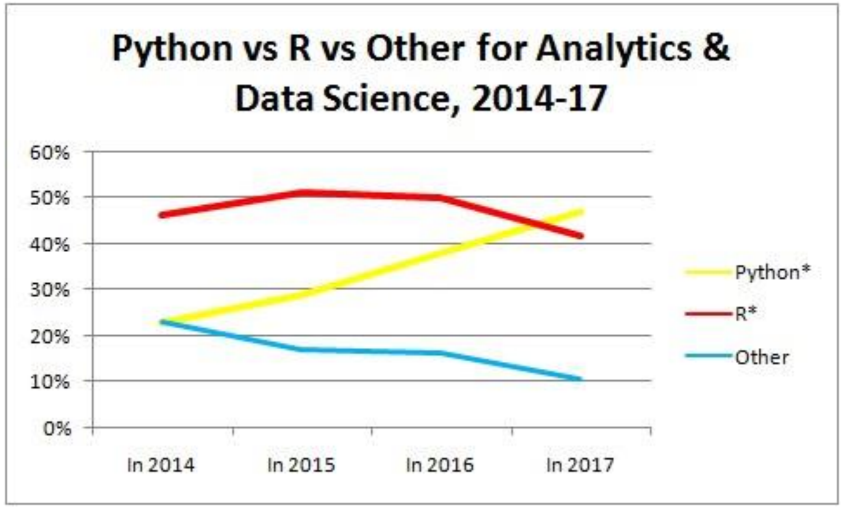


Fig. 1: Share of Python, R, Both, or Other platforms usage for Analytics, Data Science, Machine Learning, 2016 vs 2017



Next, we examine the transitions between the different platforms.

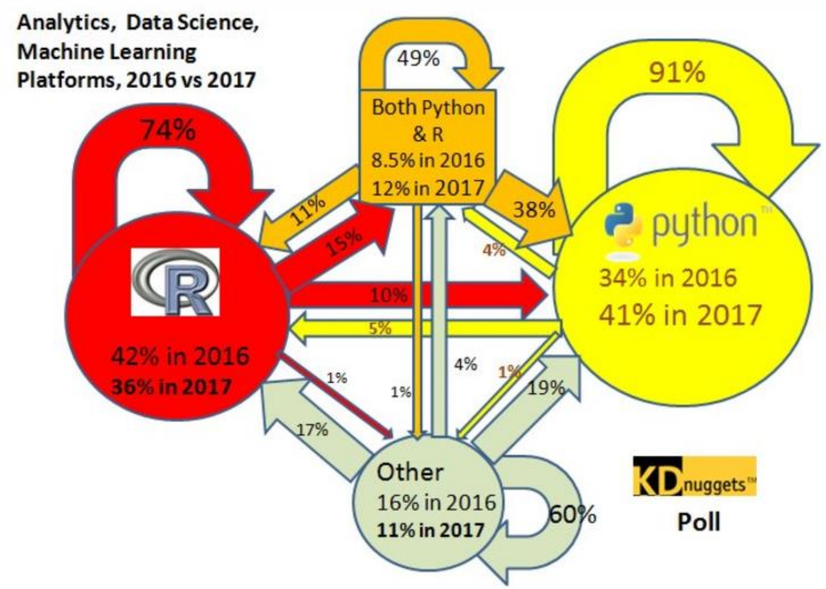
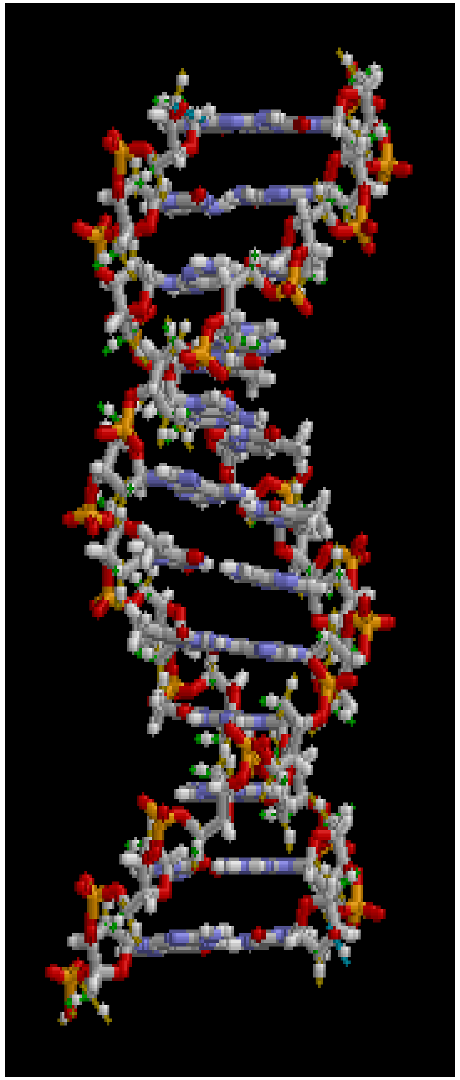


Fig. 2: Analytics, Data Science, Machine Learning Platforms Transitions between R, Python, Both, and Other from 2016 to 2017

The following Python add-ins are excellent tools to help you take advantage of multiple processors and graphical processing units (GPUs). Install them from an Anaconda prompt as follows:

```
pip install theano
pip install tensorflow
pip install keras
conda update --all
```

Genetic Algorithms



Why do we call these Genetic Algorithms?

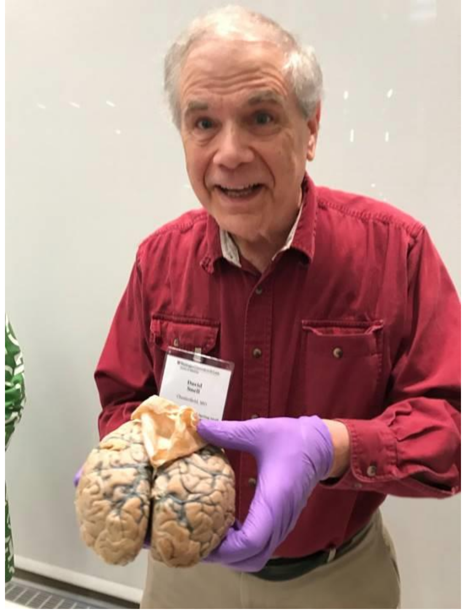
They mimic our current knowledge of genetics.

We have trillions of cells.

DNA represents a blueprint for a cell.

It is used to generate copies.

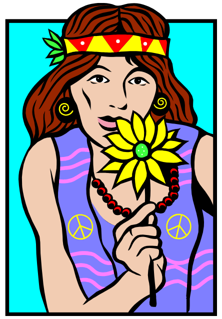
The actual process involves proteins and lots of other biological terms …



and you don't have to know them to solve problems!

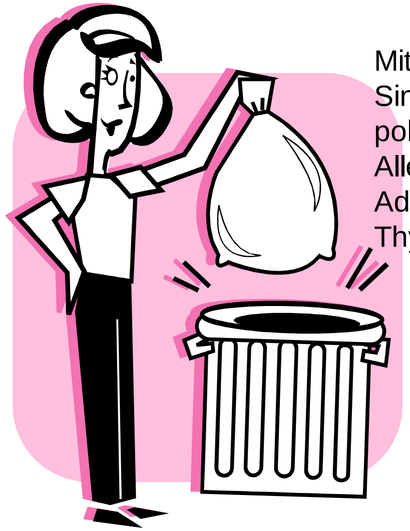
Criteria that make a problem suitable for a genetic algorithm

- **The problem involves a lot of variables - to some extent, the more variables there are, the better this technique applies.**
- **Each variable can take on potential values to produce different solutions.**
- **We can substitute a value for each of the variables and that particular combination of individual values can be thought of as a solution set.**
- **The problem can be quantified in some manner so that any two solution sets can easily be compared to see which is better.**



OMG

That is so simple!



Mitosis/meiosis
Single nucleotide polymorphisms
Alleles/ phenotypes
Adenine, Cytosine, Thymine, Guanine

Case study: Health Insurance

- A genetic algorithm running on a $300 netbook computer saved millions of dollars on a health provider network setup.
- 500 provider groups, offering 1 to 25 specialties (acupuncture to x-rays)
- If all 500 provider groups were in network, normalized cost would be 1.000
- Each group IN or OUT of network
- Potential number of solution sets was 2^{500} (a VERY big number)
- Best arrangement by experienced actuaries was cost reduction to 0.78
- Netbook got to 0.75 within an hour
- After three days it reached 0.58

MARYVILLE UNIVERSITY

Provider Network Cost Optimization

Each provider group is in (1) or out (0) of network.

					Chiropractic	Pathology	Cardiovascular Disease	Family Practice	Obstetric
5				Specialties:	Chiropractic	Pathology	Cardiovascular Disease	Family Practice	Obstetri
6				Cost	0.97	0.92	0.90	0.89	
7				Count Minimum	5	5	5	20	
8				Current Count	79	42	70	356	
9					125	62	82	597	
10		Total							
11					1	2	3	4	
12	Health System	In	Rela	Total Provider	Chiropractic	Pathology	Cardiovascular Disease	Family Practice	Obstetri
13	Provider # 1	1	0.77	M(G13:AP13)	0	23	24	64	
14	Provider # 2	1	0.90	355	1	12	26	83	
15	Provider # 3	1	0.79	287	0	0	9	66	
16	Provider # 4	0	1.13	228	0	0	0	65	
17	Provider # 5	0	0.89	216	0	0	11	67	
18	Provider # 6	0	1.36	137	0	0	0	0	
19	Provider # 7	0	1.50	129	3	17	0	10	
20	Provider # 8	0	1.32	85	0	0	0	18	
21	Provider # 9	0	1.33	38	0	0	0	0	
22	Provider # 10	0	1.08	37	0	1	0	0	
23	Provider # 11	1	1.04	35	0	0	0	0	
24	Provider # 12	0	0.73	35	0	0	0	0	
25	Provider # 13	0	1.16	34	0	0	0	0	
26	Provider # 14	0	1.32	28	0	0	0	0	
27	Provider # 15	0	1.12	27	0	0	0	4	
28	Provider # 16	1	1.03	27	0	0	0	18	
29	Provider # 17	1	0.78	26	0	0	0	0	
30	Provider # 18	0	1.22	26	0	0	0	0	
31	Provider # 19	1	1.55	25	0	0	0	1	
32	Provider # 20	1	0.84	21	0	0	0	13	

Each provider group can have multiple specialists; and has a relative cost.

500 Providers for this example; but could have thousands. Lots of specialties. Could have 2^500 (> 10^150) solution sets … might take a while by traditional methods. ☺

Provider Network Cost Optimization (continued)

Genetic Algorithm Presentation

Provider Network Fitness Function

Count of Contracts (Provider Groups) Used:	325
Included Providers (Specialists):	2,885
Relativity to Overall Network:	0.8966
Adequate Network:	Yes

Click Here to start genetic algorithm for solution set. You can modify parameters on the Parameters sheet.

Specialty	Available Providers	Required Providers	Selected Providers	Requirement Met	Relativity	Specialty Weight
Hospital	16	5	11	Yes	0.89	47.1%
Family Practice	597	20	438	Yes	0.90	7.7%
Physical Therapy	506	5	243	Yes	1.00	3.9%
Internal Medicine	376	20	296	Yes	0.89	3.8%
Obstetrics/Gynecology	277	5	195	Yes	0.88	3.8%
Pediatrics	351	5	249	Yes	0.95	3.4%
Orthopedic Surgery	147	5	100	Yes	0.88	3.2%
Hematology /Oncology	97	5	58	Yes	0.86	2.8%
Chiropractic	125	5	87	Yes	0.98	2.7%
Diagnostic Radiology	174	5	101	Yes	0.87	2.5%
Dermatology	61	5	47	Yes	0.81	2.1%
Ophthalmology	120	5	111	Yes	0.86	1.3%
Otolaryngology	52	5	45	Yes	0.82	1.3%
Gastroenterology	40	5	34	Yes	0.89	1.2%
Pathology	62	5	41	Yes	0.90	1.1%
Podiatry	44	5	32	Yes	1.01	1.0%
Acupuncturist	65	5	39	Yes	0.94	0.9%
Urology	44	5	32	Yes	0.93	0.9%
General Surgery	65	5	46	Yes	0.84	0.8%
Rheumatology	21	5	16	Yes	0.86	0.8%
Neurology	94	5	86	Yes	0.91	0.8%

Instructions | Summary | Provider_List | DNA | Parameters

Each specialty area must have adequate coverage.

Some problems just don't fit well into classical methods of solution:

Assume you have three equations:

- $y_1 = a * e * g + h + d^a$
- $y_2 = |h|! - |d|!$
- $y_3 = ((\sin(a)) + b) * \log(b + c))$
  $+ \cos(\min(c, d)) * (e - f + g * h)$

Find a combination of ***a, b, c, d, e, f, g, h*** such that the standard deviation of y_1, y_2, and y_3 is minimized.

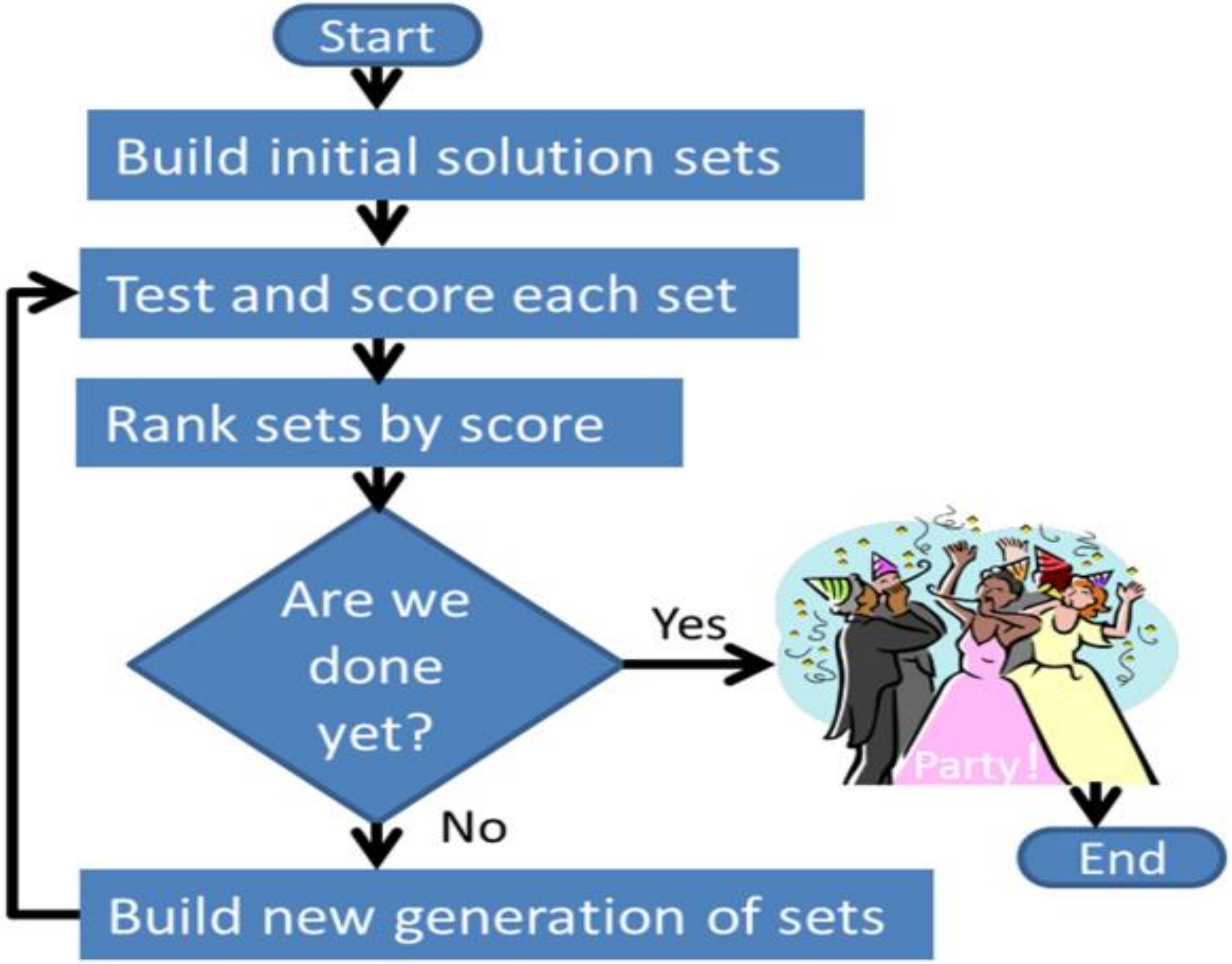
Oh yeah!
We are math folks, so this might be too easy by itself!

Let's add some constraints to make it more interesting!

a has to be an integer from 1 to 10
b is a real number from 0 to 15
c is a real number from 1 to 3
d is a real number from 0.5 to 7
e is a real number from -10 to 50
f is an **even** integer from -20 to 40
g is a real number from 0 to 18
h is a real number from 3 to 12

Click for Live Demo

How to build a genetic algorithm



VBA example code

from elsewhere: elites = 20
setsPerGeneration = 100
parentPool = 40
solutionSets is a
2-dimensional array 30 by100

```
Private Sub AddTheChildren()
Dim parent As Integer, var As Long, child As Integer, children As
Integer
  children = setsPerGeneration - elites
    For child = 1 To children
      For var = 1 To setLength
        parent = Int(parentPool * Rnd()) + 1
        solutionSets(var, elites + child)
                                      = solutionSets(var, parent)
      Next var
    Next child
End Sub 'AddTheChildren
```

(80 = 100 – 20)

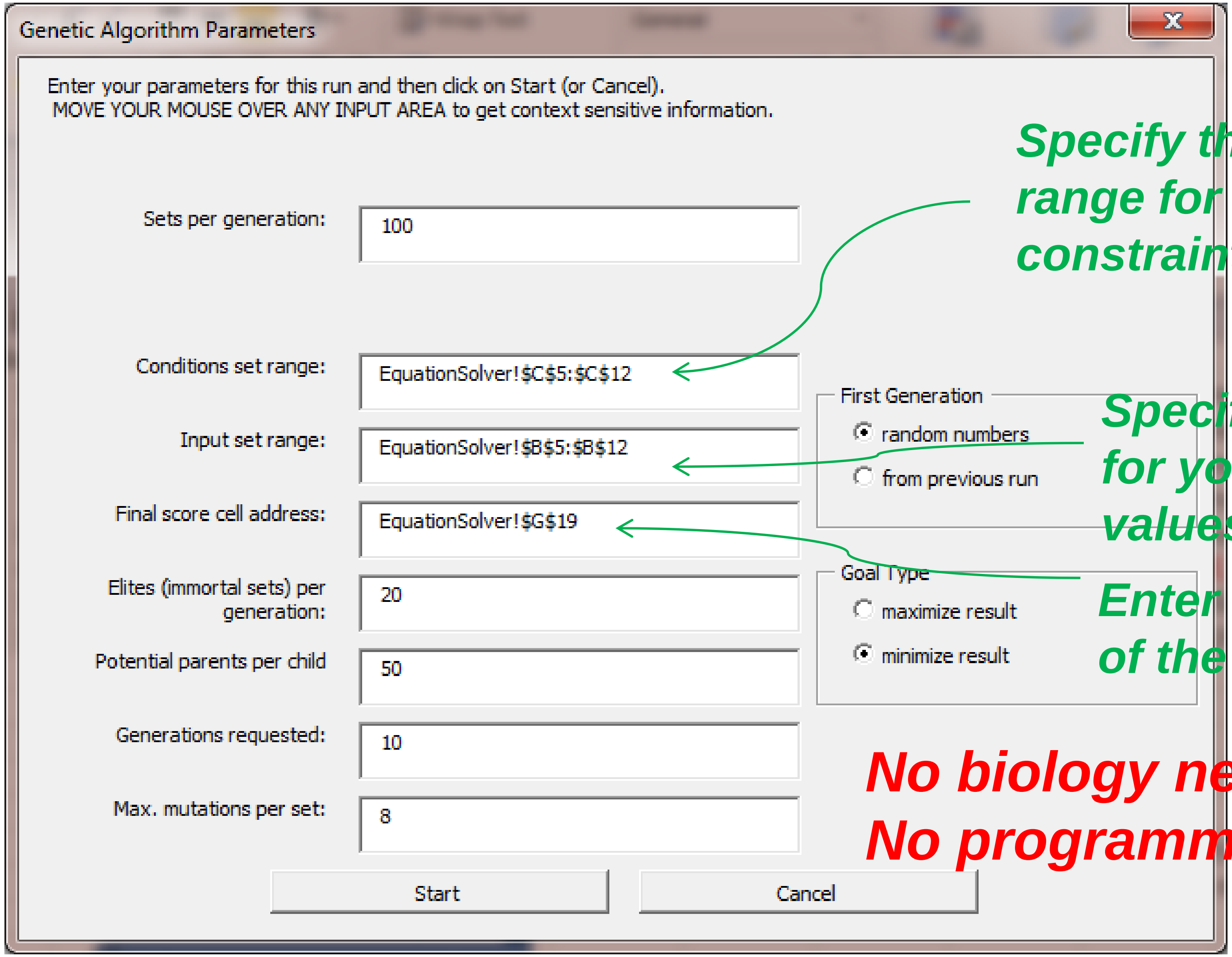
(start with child set 1)

(1 to 30 if 30 variables per set)

(e.g. parent 5 wins for variable 17 for child 1)

(so set variable 17 in new solution set 21 = 20 +1 to the value from variable 17 in old solution set 5)

Input Screen for *FREE* workbook



Specify the sheet and range for your constraint criteria.

Specify the range for your variable values

Enter the location of the final score

No biology needed!

No programming needed!

MARYVILLE UNIVERSITY

Overview of Predictive Analytics Techniques

Frequent Itemsets

Checking Rx history to infer 'other' drugs taken by applicant

Jeff Heaton
Data Scientist
RGA Reinsurance Company

Dave Snell, ASA, MAAA
Technology Evangelist
RGA Reinsurance Company

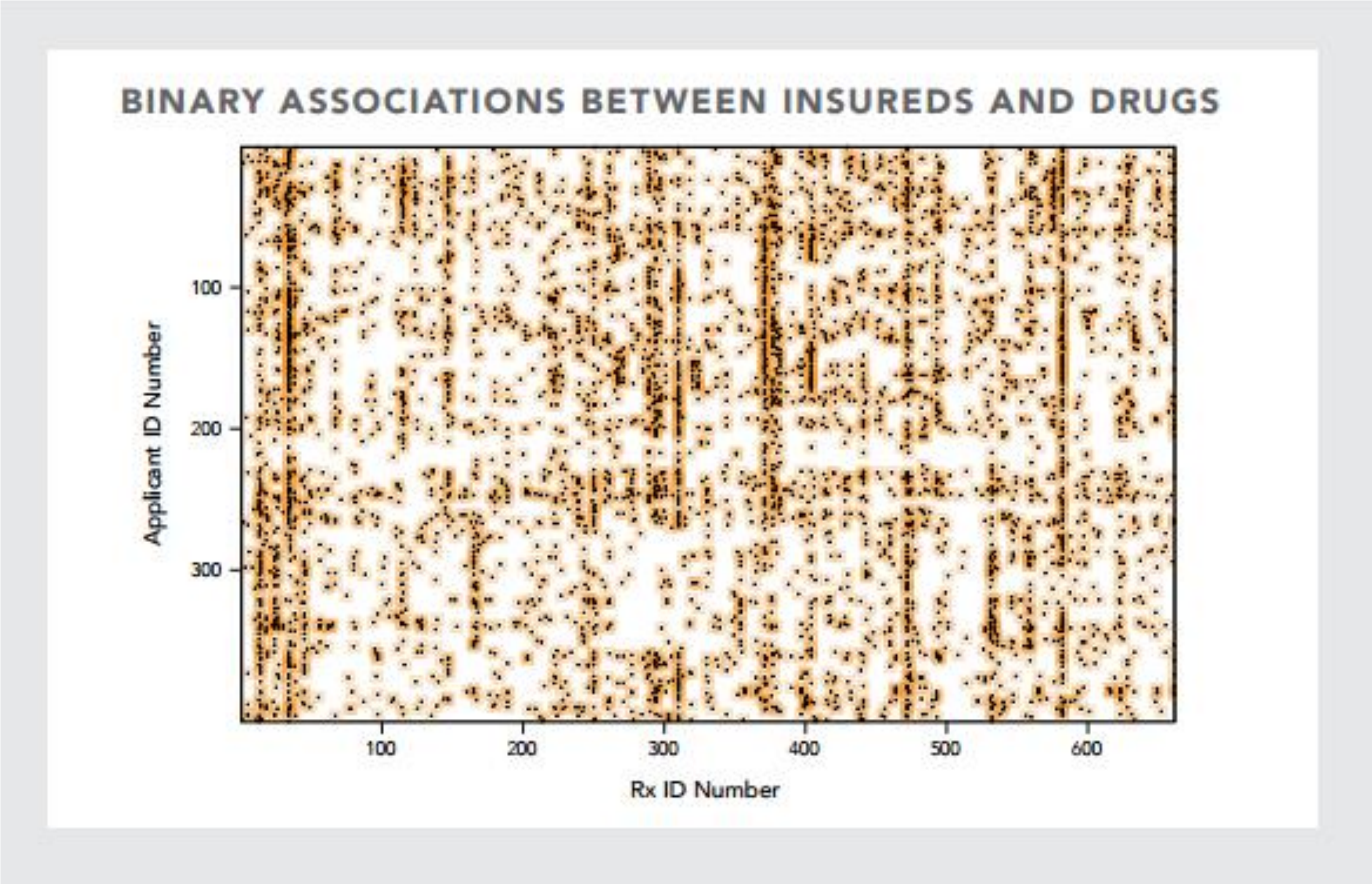
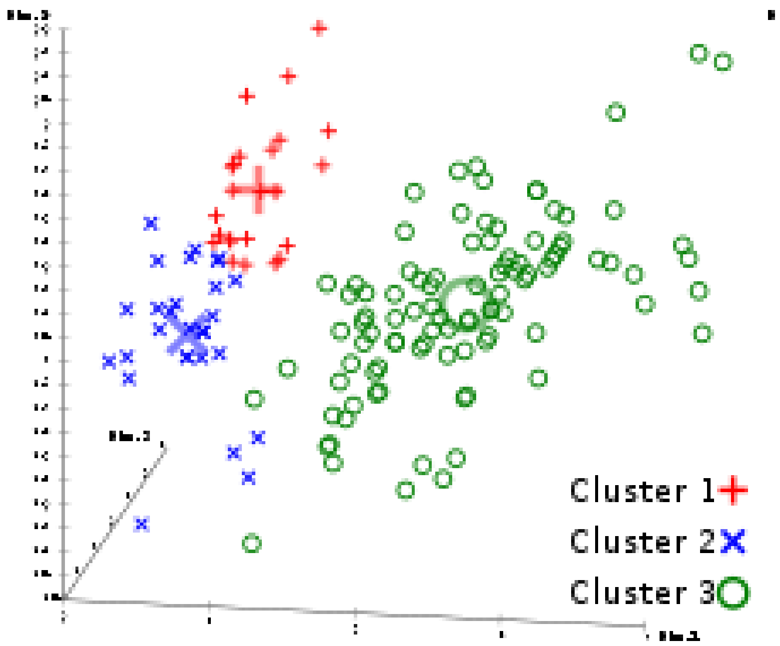


Image from paper by authors

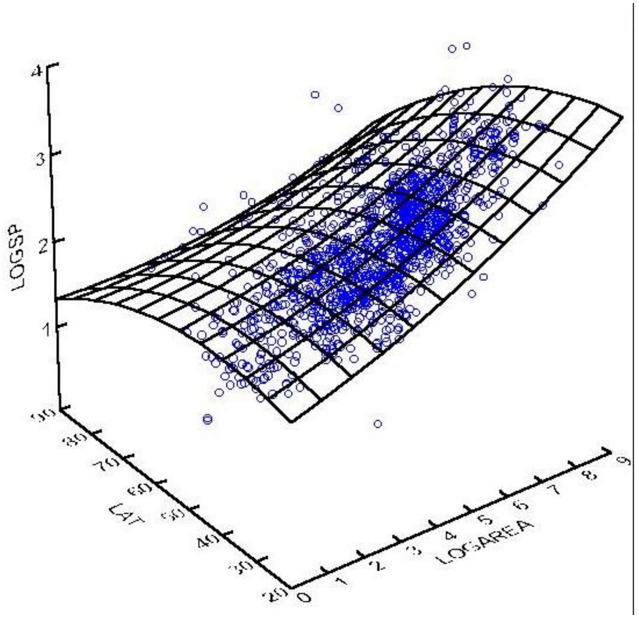
Utilizing Frequent Itemsets to Model Highly Sparse Underwriting Data – Society of Actuaries - Predictive Analytics 2014 Call For Articles

Overview of Predictive Analytics Techniques

K-Means Clusters



Multi-variate Regression



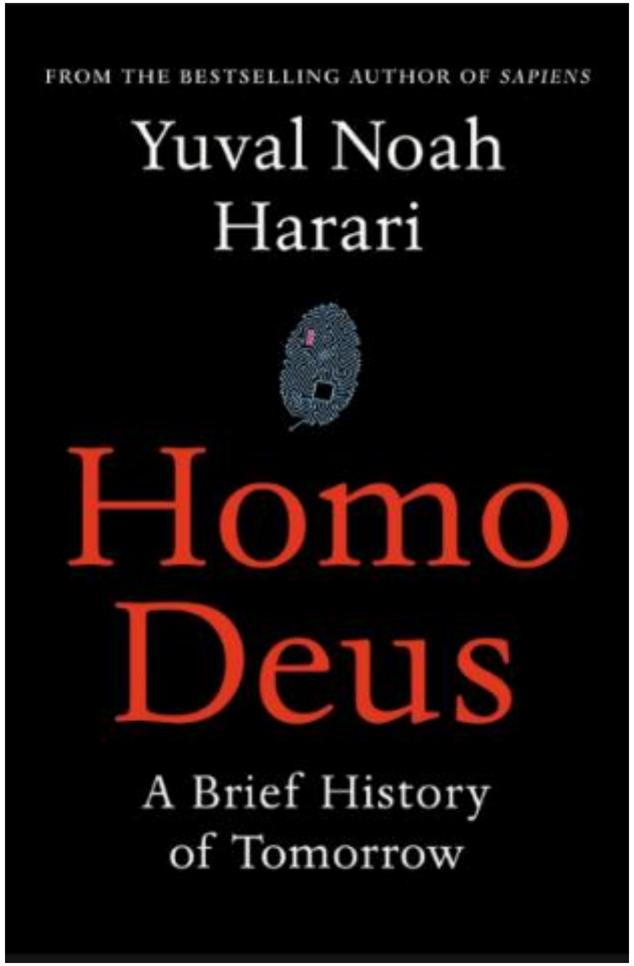
Images from Wikipedia and designated in public domain

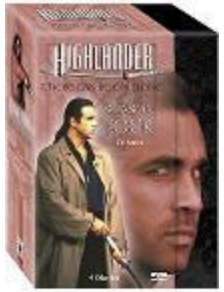
For the first time in history, more people die today from eating too much than from eating too little;

more people die from old age than from infectious diseases;

more people commit suicide than are killed by soldiers, terrorists and criminals combined.

A few serious scholars suggest that by 2050, some humans will become a-mortal (not immortal, because they could still die of some accident, but a-mortal, meaning that in the absence of fatal trauma their lives could be extended indefinitely).





This Photo by

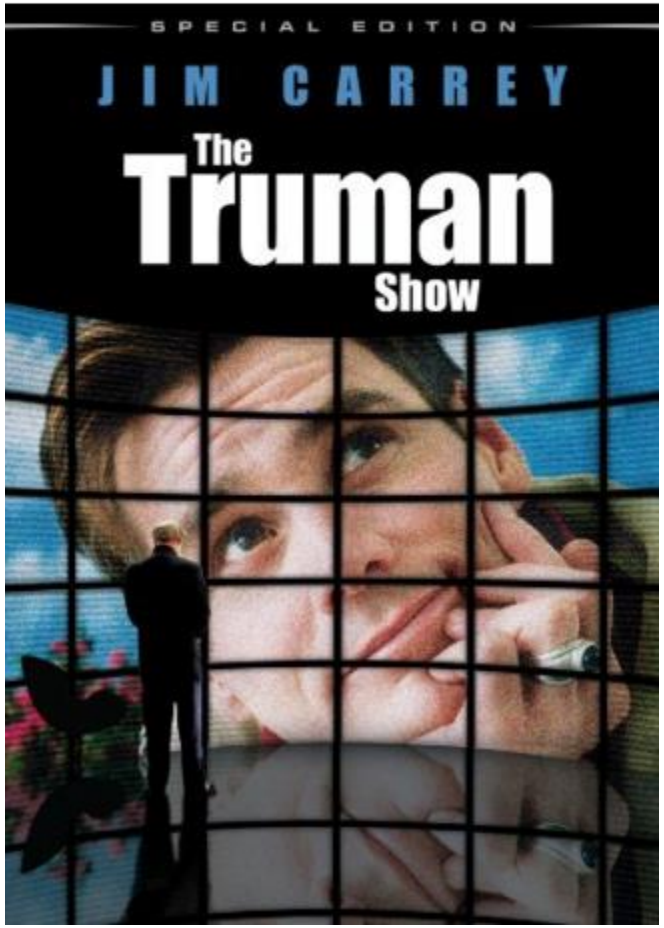
How are Big Data and predictive analytics changing healthcare?

The Truman Show was just the Beginning!

Genome
Phenome
Physiome
Anatome
Transcriptome
Proteome
Metabolome
Microbiome
Epigenome
Exposome

A Panomic perspective!

Try
http://www.wolframalpha.com/facebook/
but be very afraid!



How are Big Data, predictive analytics, and machine learning changing healthcare?

"In every other industry, technology drives down costs and consumers are considered perfectly capable of making decisions for themselves." –David Goldberg

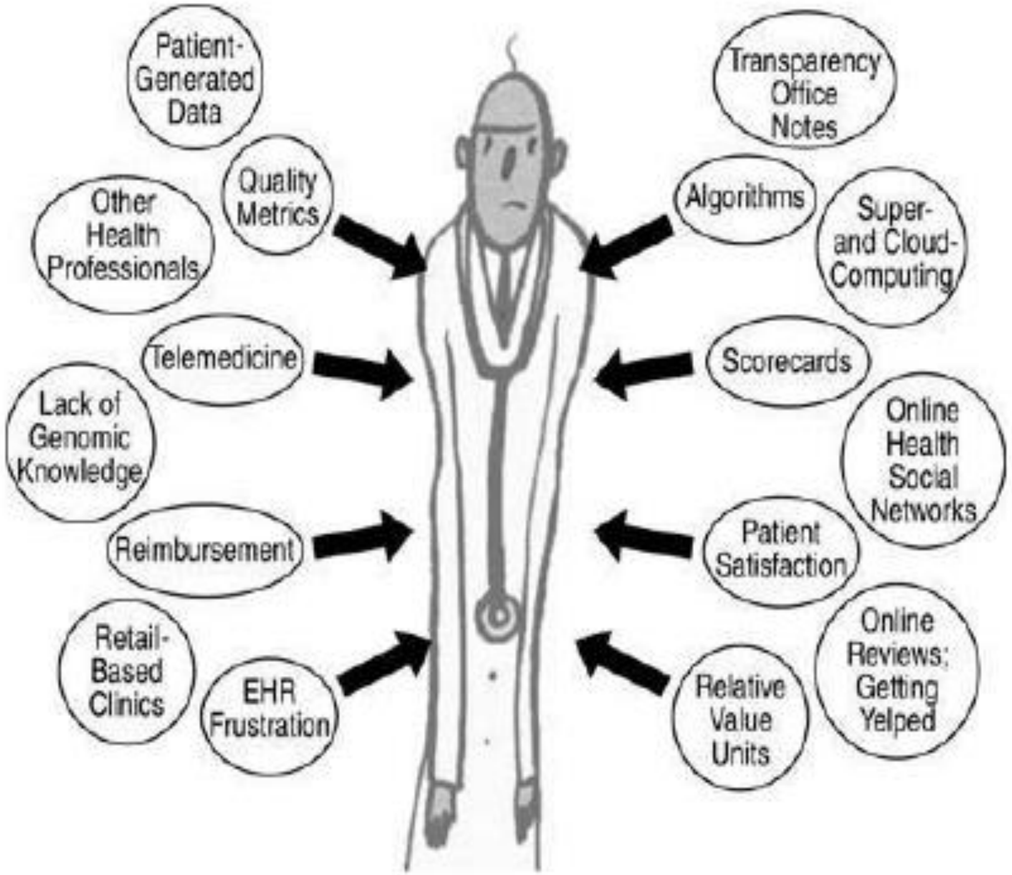
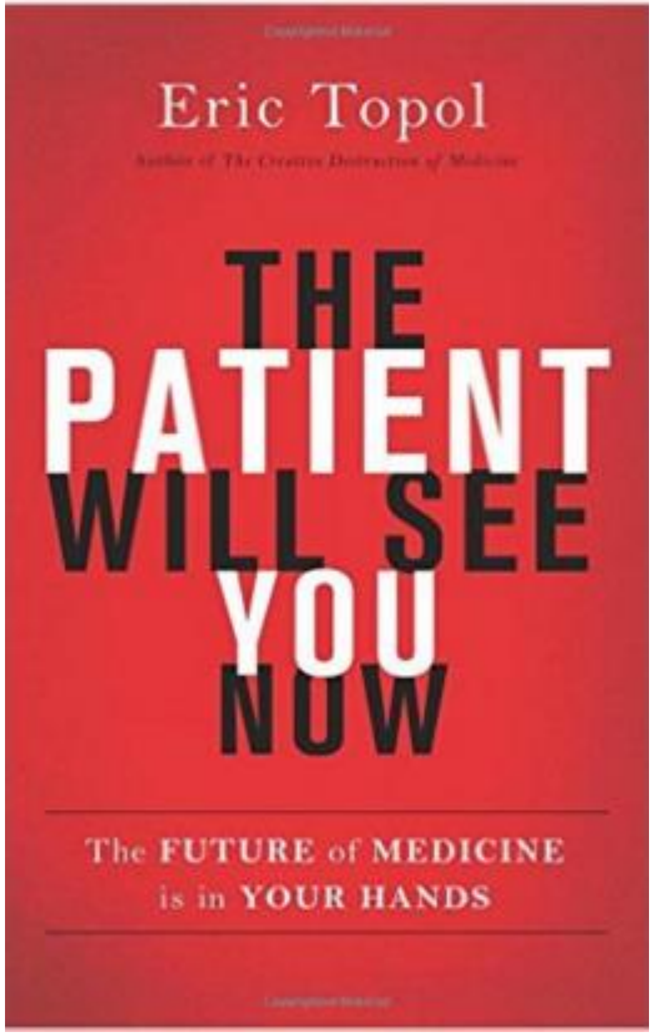


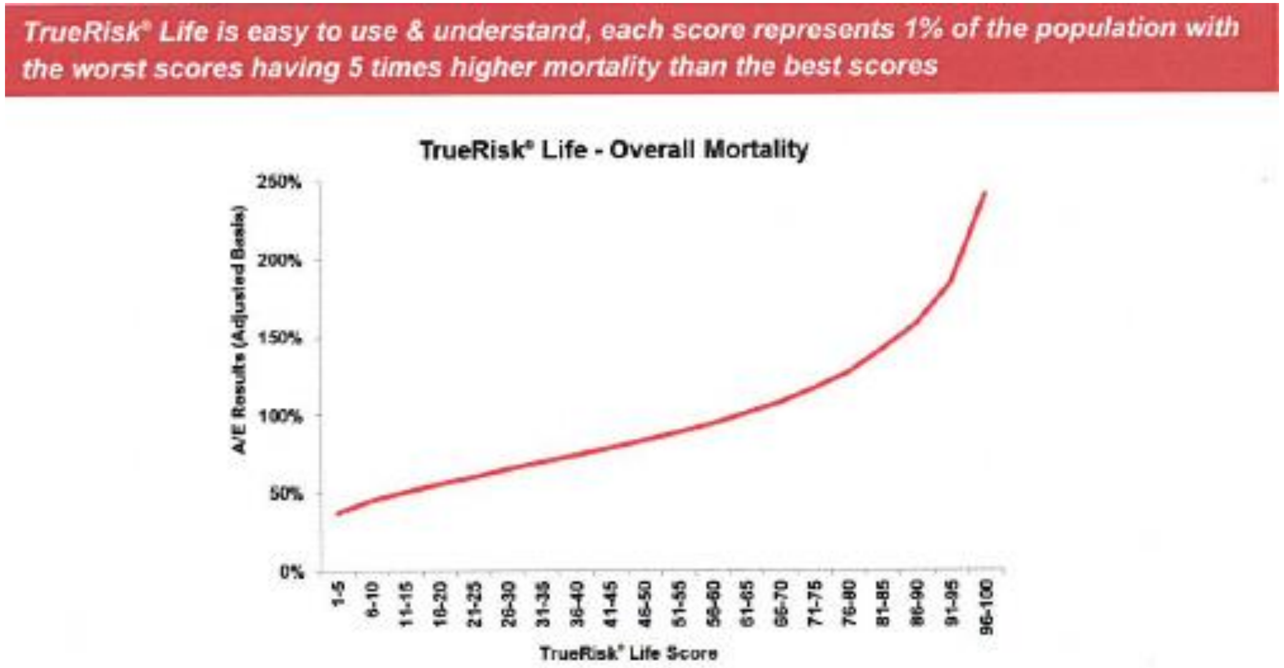
FIGURE 9.4: Doctors are getting squeezed like never before.



Images used with permission from Eric Topol, M.D.

How is Machine Learning already changing insurance underwriting?

- A short history lesson
- The Attending Physician's Statement (APS) – "Gold Standard?" "fool's gold?"
- Pharmaceutical (Rx) databases
- TransUnion TrueRisk® Life
  - http://www.rgare.com/knowledgecenter/Documents/RGAWebcastCreditBasedSolutions.pdf



Top 5% (worst scores) had six times the early policy lapses, (years 1 and 2) and five times the mortality (years 1 – 12) of the lowest (best) 5%.

(lessons learned – people do not follow strictly analytic models)

- Logical Rule: People are rational and make rational decisions
- Logical Rule: Accuracy is more important than marketing hype
- Logical Rule: Everyone acts in a manner that will maximize their own self-interest
- Logical Rule: The work of science is to Substitute facts for appearances and demonstrations for impressions – John Ruskin

(Motto of Society of Actuaries)

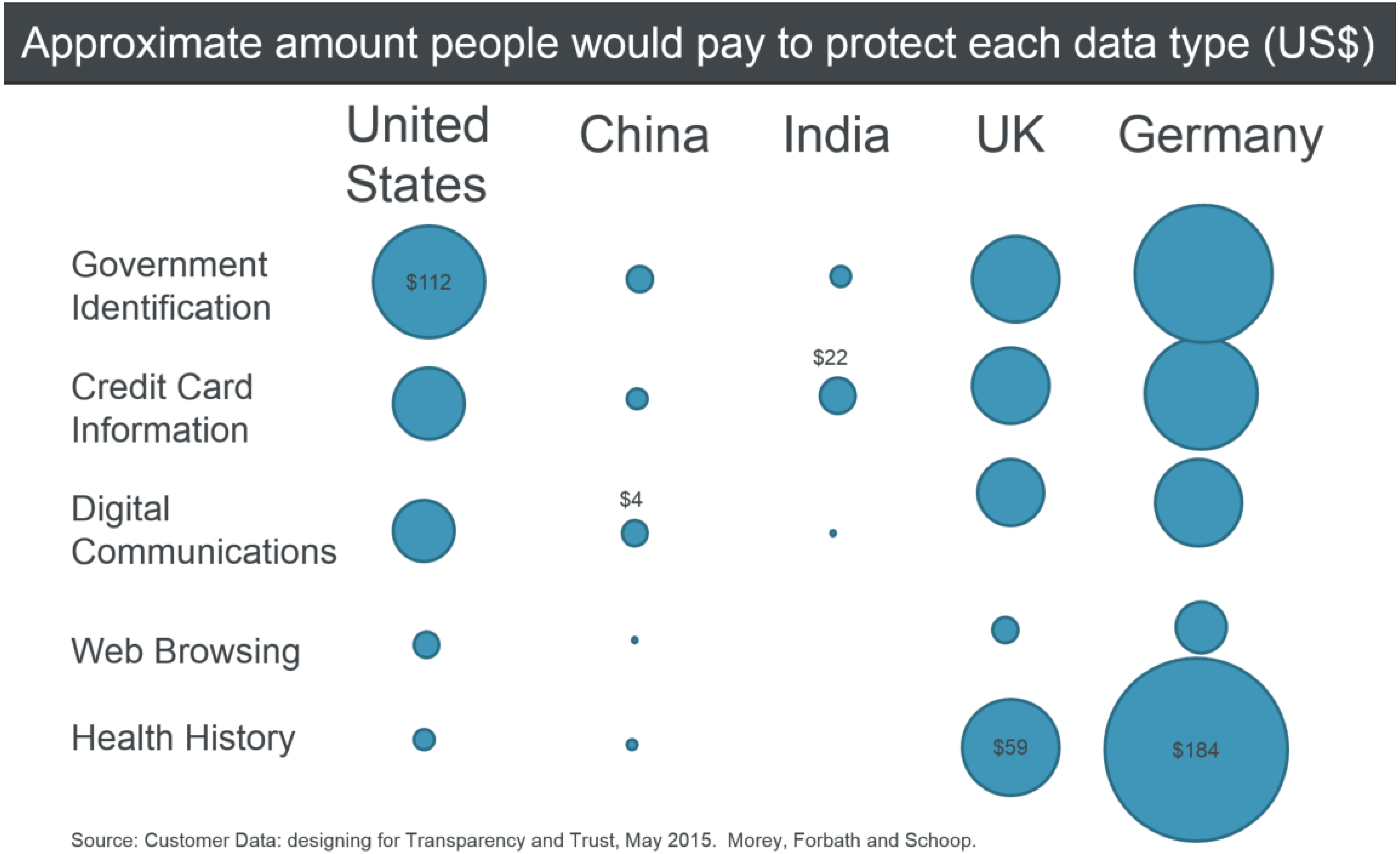
MARYVILLE UNIVERSITY

Analytic Models Fail When They Ignore Behavioral Economics

http://www.jasonheadley.com/INATN.html

(shown with permission – if you like this video, consider putting a tip in the Vimeo tip jar: https://vimeo.com/66753575.)

Consider Cultural Differences A 'Home Run' may not 'Hit a Six'



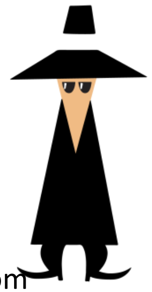
Source: Customer Data: designing for Transparency and Trust, May 2015. Morey, Forbath and Schoop.

Original source included other types of data, all with low amounts (e.g. search history, location, purchase history, contact information)

How are they already changing insurance underwriting? (lessons learned)

- Target uses predictive modeling to determine a teenage girl is pregnant
    - They send her lots of coupons for baby products (Forbes: 16-Feb-2012)
    - Their analytics are correct, but it backfires on them.
- Progressive Insurance Company
    - The first major(?) insurance company to employ predictive analytics in a big way.
    - Huge increase in market share.
    - Huger increase in bottom line profit.
    - What a wonderful endorsement of predictive analytic models!
    - But that was only part of the picture.

Behavioral Economics was the *secret sauce* that made it palatable to the public.



RGA's Experience in Predictive Analytics in Life and Health insurance

Sales, Marketing & Distribution

Propensity to Buy: Use client & sales data to identify best leads for marketing efforts. Project in Taiwan.

Agent Quality Assessment: Use policyholder & claims data to determine which agents add most value to profitability. Projects in US and UK.

Policy & Claims Management

Experience Analysis: Use multivariate model to understand the true drivers of experience. Projects in US, UK and South Africa.

In Force Retention: Use customer & lapse data to determine which policies are most likely to lapse in order to develop retention strategies. Projects in US and UK.

Fraud Detection: Use customer and claims data to determine which claims are most likely to be fraudulent and focus forensic efforts on them. Project in India.

Underwriting Improvements

Predictive UW / Cross-Sell: Use customer & UW decision data to model mortality risks. Most successful if insurance data is combined with other data sources e.g. Banking, P&C, Retail etc. Projects in HK & SEA, Japan, Australia, US.

Claims Experience/ Up-Sell: Use customer & claims data to segment or automatically underwrite in-force policyholders for new or existing products. Project in Japan.

Guideline Refinement: Use underwriting and claims data to determine optimum UW requirements e.g. adding or removing questions; optimize non-medical limits etc. Project in China.

UW Investigation: Use customer and UW investigation data to determine which cases are most likely to require further investigation. Project in Korea.

Case Prioritization/Triage: Use customer and placement data to determine which cases are most likely to be successfully placed and focus resources on them. Project in the US.

Multi-line Predictive UW Cross-sell

A multi-line insurance company with a large P&C customer base expressed a strong desire to increase the sales penetration of their life product, while streamlining the underwriting process.

Objectives

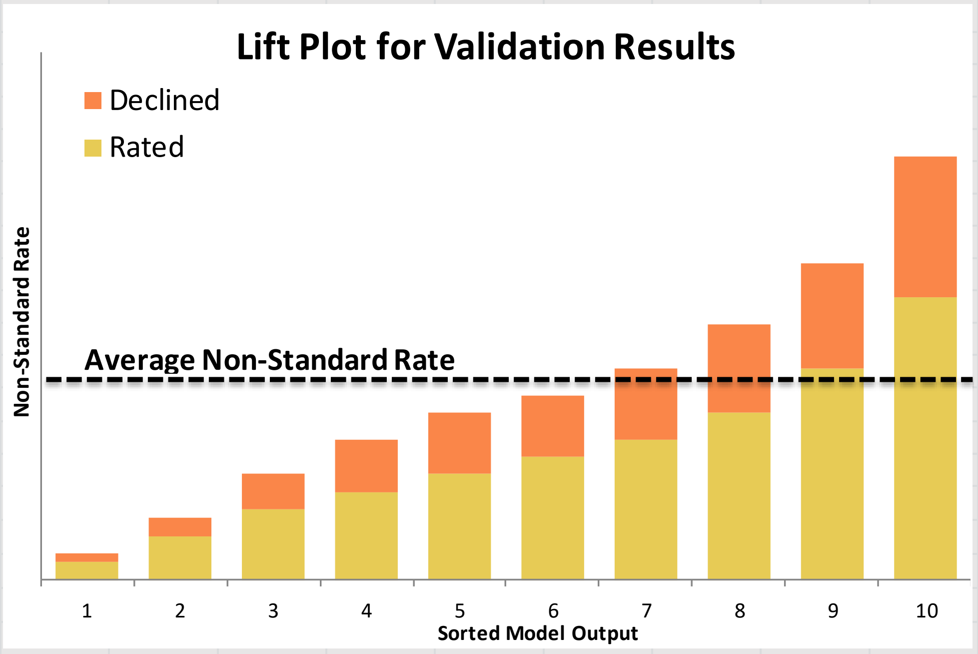
- Increase life product penetration
- Improve customer experience for best risks with a simplified UW & sales process
- Improve persistency of P&C customers as a result of a deeper client relationship

Data

- Combined Life + P&C data (at time of UW), including rated & declined; enhanced with rich 3rd party data
- One model used to predict smoker status and 2nd model used to predict being preferred
- At least two dozen variables used for each model, e.g. age, gender, auto violation points, vehicle maker, etc.

Business Application & Lift Plot

- >20% of current P&C policyholder selected by model
- Only 1 UW question asked for pre-selected customers for life product at standard rate to replace medical UW



Bancassurance Predictive Underwriting

A bank with a large customer base expressed a strong desire to increase the sales penetration of their life product, while streamlining the underwriting process.

Objectives

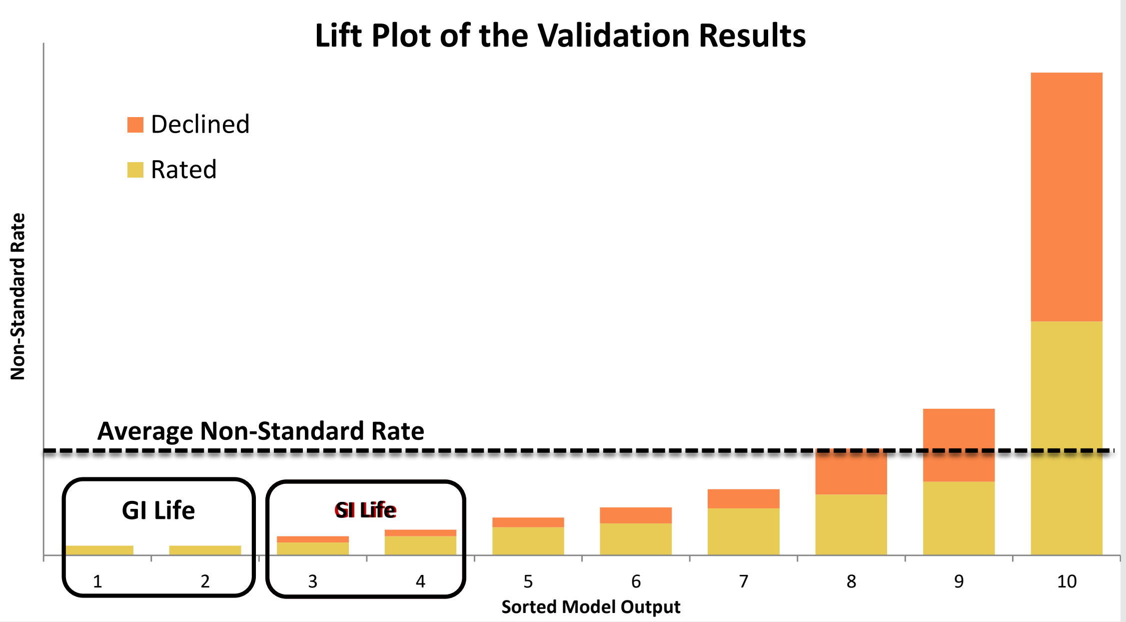
- To have a simplified underwriting and sales process with high take-up for the best risks
- To reduce acquisition cost
- To improve financial performance

Data

- Two data sources combined
  - Underwriting data at the time of Issue
  - The bank's financial database
- About 80 variables available for modeling
  - For ex. Demographic Info, Bank & Insurance Product Info and Bank Transactions

Business Application & Lift Plot

- 11 statistically significant variables in model,
  - Age, gender, branch, AUM, customer segment
- No underwriting questions for the best 20% risks; next best 20% for simplified issue with very few UW questions



MARYVILLE UNIVERSITY

Medical Product Upsell

An insurance company has a sizeable medical product customer base and would like to up-sell additional medical coverage to in-force policyholders with significantly reduced underwriting.

Objectives

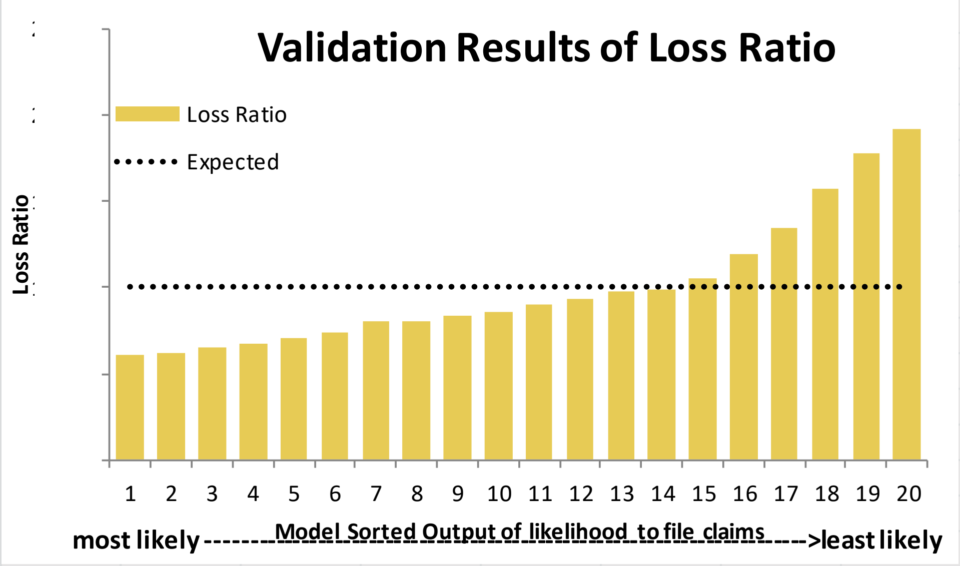
- To increase sales by upselling new PHI products to in-force customer base
- To increase take-up by reducing underwriting requirements for the best risks

Data

- Large data set with detailed policyholder & claims information (more than 3m base policies & claims, more than 4m riders)
- Modelled total claim cost using wide range of rating factors & compared to pricing to identify low risk policyholders

Business Application & Validation

- Additional variables identified beyond age/gender, including location, income, occupation, rider count, etc.
- Upsell the new product to 50% of current in-force customers with one UW question on pre-approval basis
- Significantly simplify sale process for customers & agents



Cancer Product Upsell

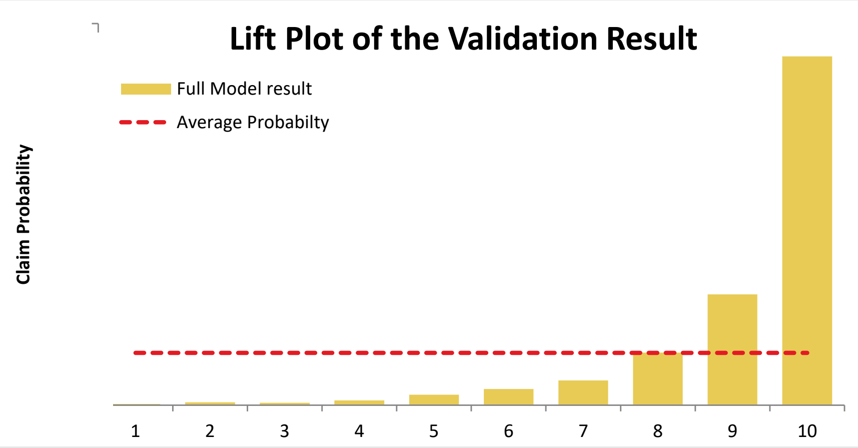
An insurance company has a sizeable cancer product customer base and would like to up-sell a new life & cancer combined product to in-force policyholders with significantly reduced underwriting.

Objectives

- Find the best risks in the in-force customer base to sell the new product to
- Improve claims experience for this new product

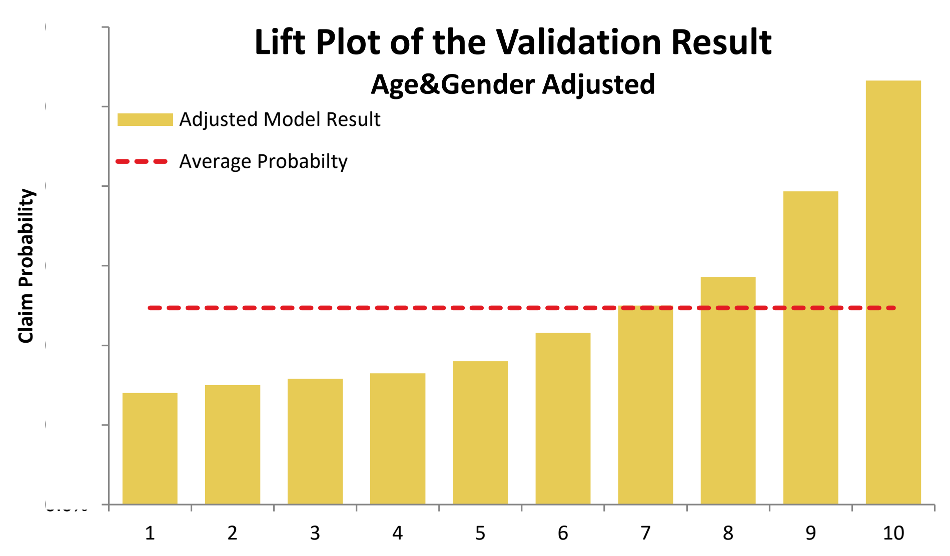
Data

- Received a very comprehensive data set to understand the true drivers of experience



Business Application & Validation

- Select best 50% of current customer for the new product with expected future improved claim experience by 40%
- Apply tightened UW for the worst 20% which has an average incidence rate 80% higher than the average
- Effective model with predictive power on claims beyond the current rating factors of age/gender:



MARYVILLE
UNIVERSITY

Non-Medical Limit Risk Segmentation

A life insurer would like to adjust their non-medical limits based on real claim experience. Limits to vary according to bespoke risk segmentations instead of a single limit for all risks.

Objectives

- Determine optimal non-medical limits that should be used for different customer segments
- Streamline the underwriting process
- Identify low risks for up-sell or cross-sell campaigns
- Find true drivers of experience to improve business decisions

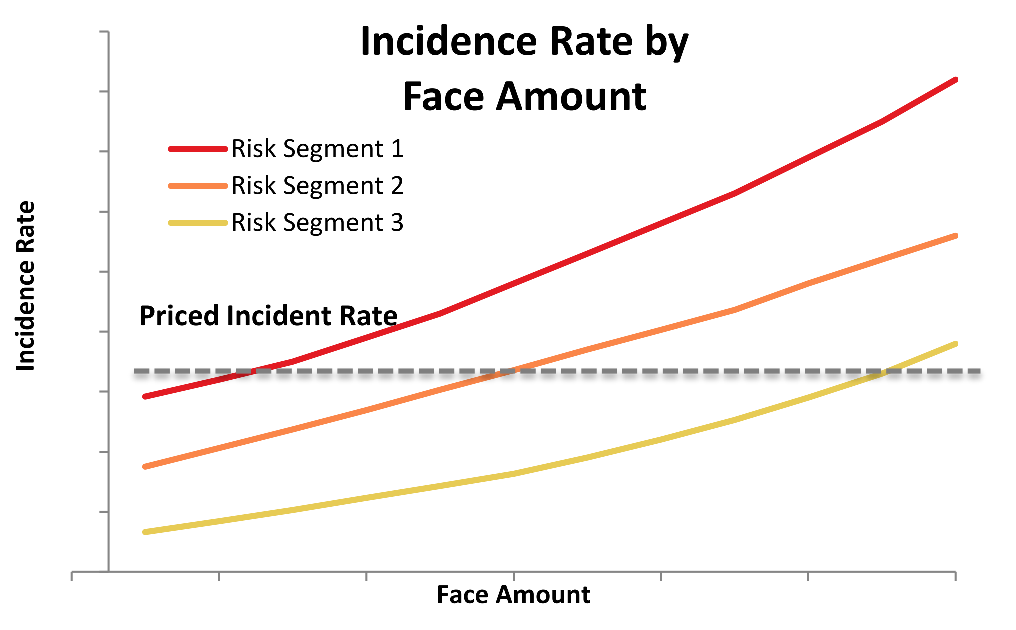
Data

- Insurance company policyholder, agent & underwriting data

Data Source	In-force + claims
Study Period	5 years
Product	Life & Accelerated CI
Total Exposure	Around 7m life years
Total Claims	Around 10,000

Business Application & Results

- Justify the requirement of non-medical limit to mitigate risk in the target markets
- High sales volume with high non-medical limit at controlled risk level for good risks



Claims Fraud Detection

An insurance company was interested in a consistent and effective fraud detection procedure which optimized the use of limited investigation resources.

Objectives

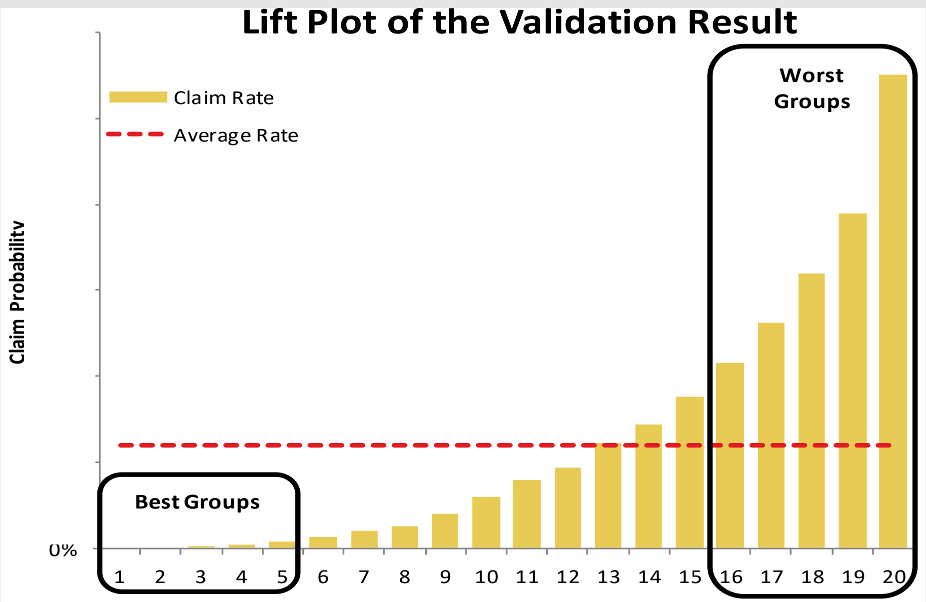
- Determine how likely an incoming claim is a fraudulent case and take appropriate action
- Make the best use of limited claim investigation resources

Data

- Use claims experience data that have been investigated with known results
- Combined customer demographic information, policy information, and claim background

Business Application & Validation

- No need to allocate any resource for the best 25% claims, while put vigorous investigation on the worst 25%
- Gain insights on the driving factors of fraud cases to incorporate into pricing bases



How Can an Actuarial Model Fail?

Formula for Ruin:

Take two Nobel prize winning economists.

Add their highly sophisticated mathematical model – even more sophisticated than the one that became part of the actuarial study notes.

Result: $3.625 Billion Bailout!

Lesson: Very intelligent people can make huge mistakes when they ignore the likelihood of illogical actions

An Insurance Example

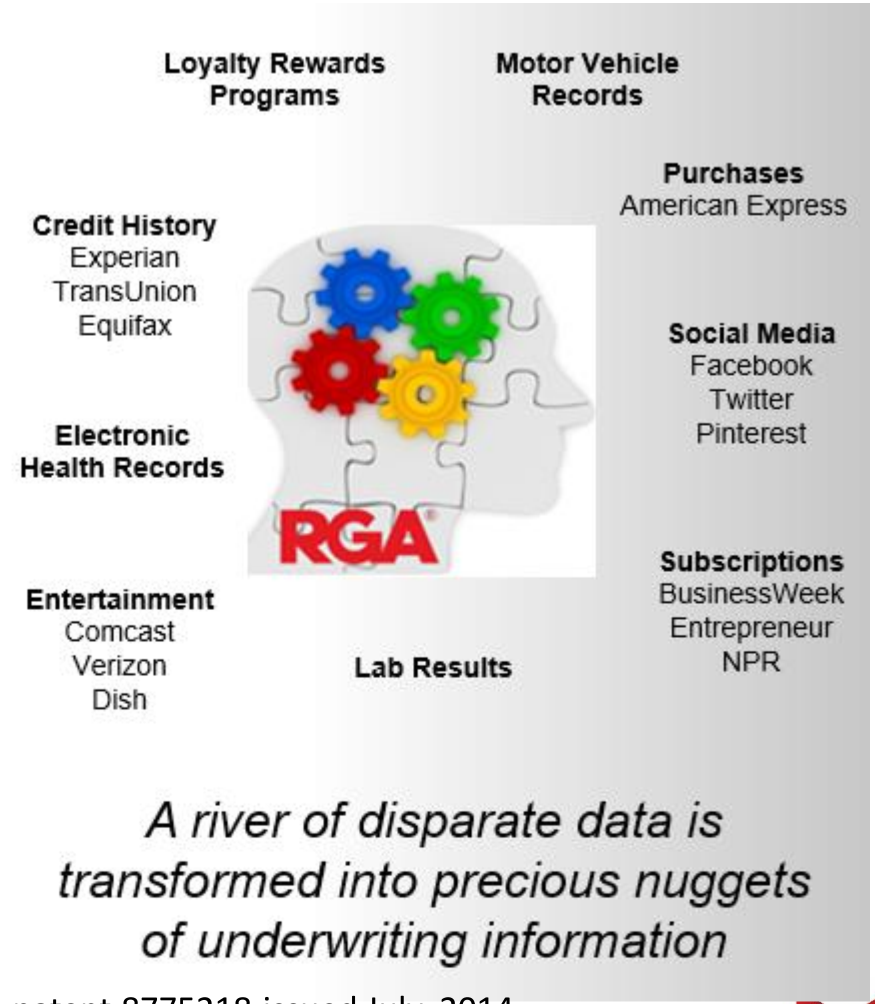
- When is 6 billion dollars more important than 30 billion dollars?
- When it causes the 30 billion dollar asset company to go into receivership!

A conditional tail expectation (CTE)
may be necessary but not sufficient!

How will AI and machine learning dramatically change future life insurance underwriting?

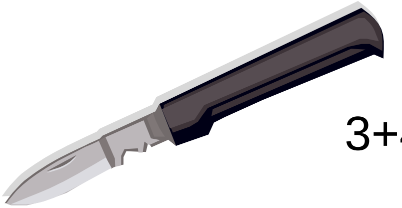
The internet of things will know more about you than any personal doctor could ever hope to know about you.

- Wearables; watches, shirts, socks, etc.
- Embeddables: pills, nanobots, labs in your bloodstream
- Appliances: smart fridge, ‘lav’ results, Kindle reading, movies and shows watched
- Consumables: the telltale hamburger, bragging broccoli
- These go beyond Big Brother’s wildest dreams!

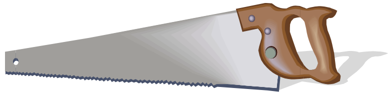


RGA patent 8775218 issued July, 2014
Transforming data *for rendering* an insurability decision

New Tools Require New Skills

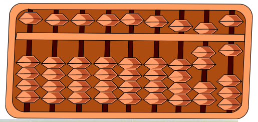


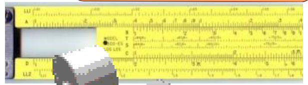
3+4=4+3

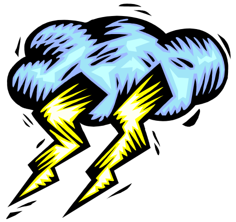


3-4≠4-3

Find this video on YouTube via search term ChainsawAlton (one word).



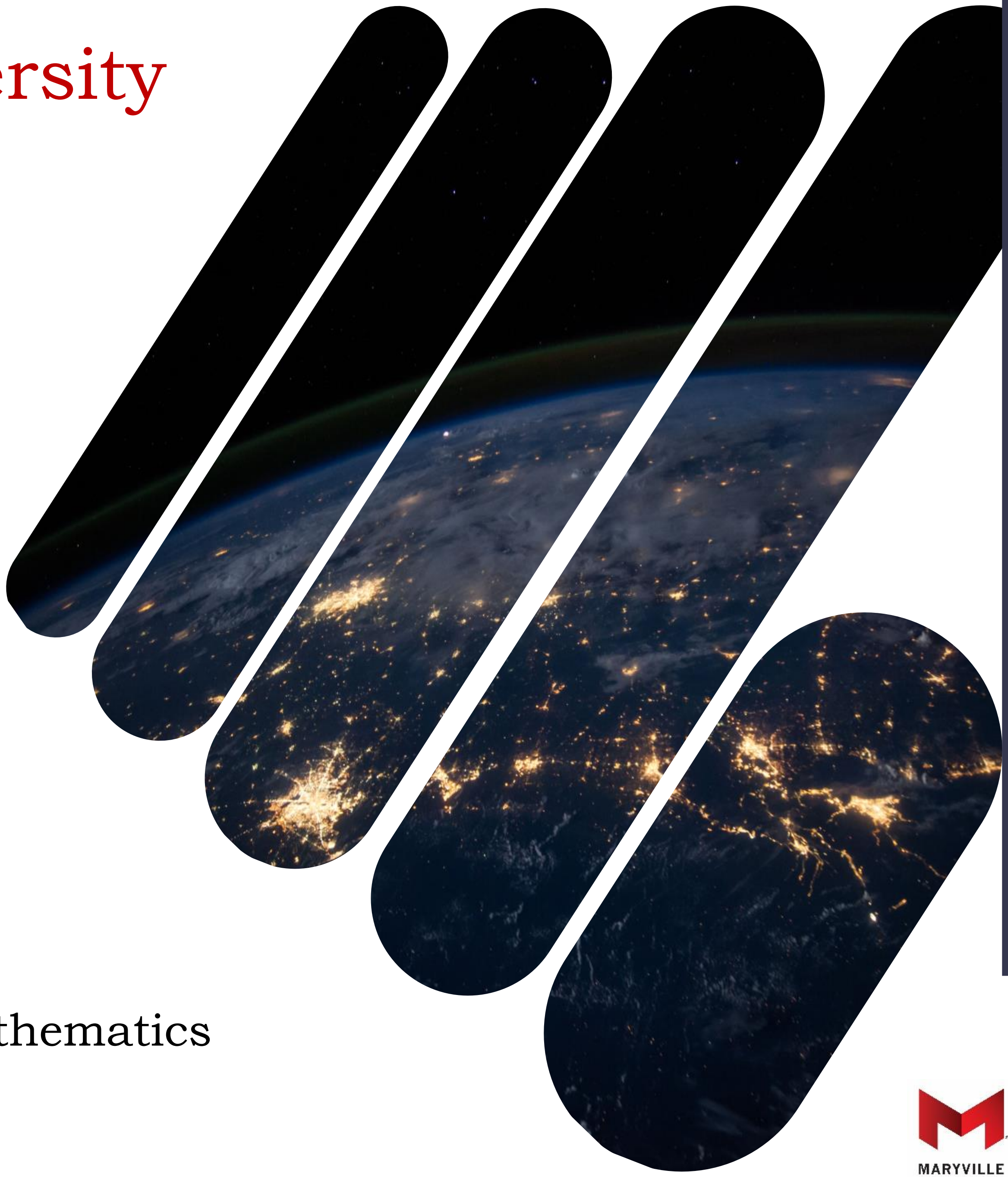




Complexity
Sciences

Data Science at Maryville University

MS in Data Science
BS in Data Science
Minor in Data Science



Courses

- Machine Learning
- Big Data Analysis
- Introduction to R
- Python
- SQL

- Everyday Data
- Introduction to R
- Predictive Modeling
- Deep Learning
- Statistical Modeling

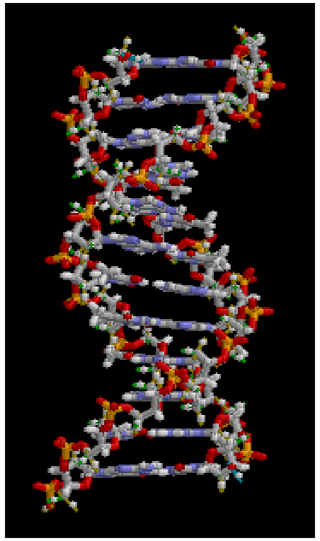
<u>Jennifer Yukna</u>
Assistant Dean of Sciences & Mathematics
jyukna@Maryville.edu
314-529-6858

<u>Guangwei Fan</u>
Director of Actuarial Science, Data Science, & Mathematics
gfan@Maryville.edu
314-529-9427

MARYVILLE UNIVERSITY

Artificial Intelligence and Machine Learning: What are they? How can we adapt to them?

Introduction to Machine Learning
St. Louis Actuaries Club 26-Jun-2018



Dave Snell, FALU, FLMI, ASA, MAAA, CLU, ChFC, ARA, ACS, MCP
technology evangelist, SnellActuarialConsulting - dave@ActuariesAndTechnology.com

MARYVILLE UNIVERSITY