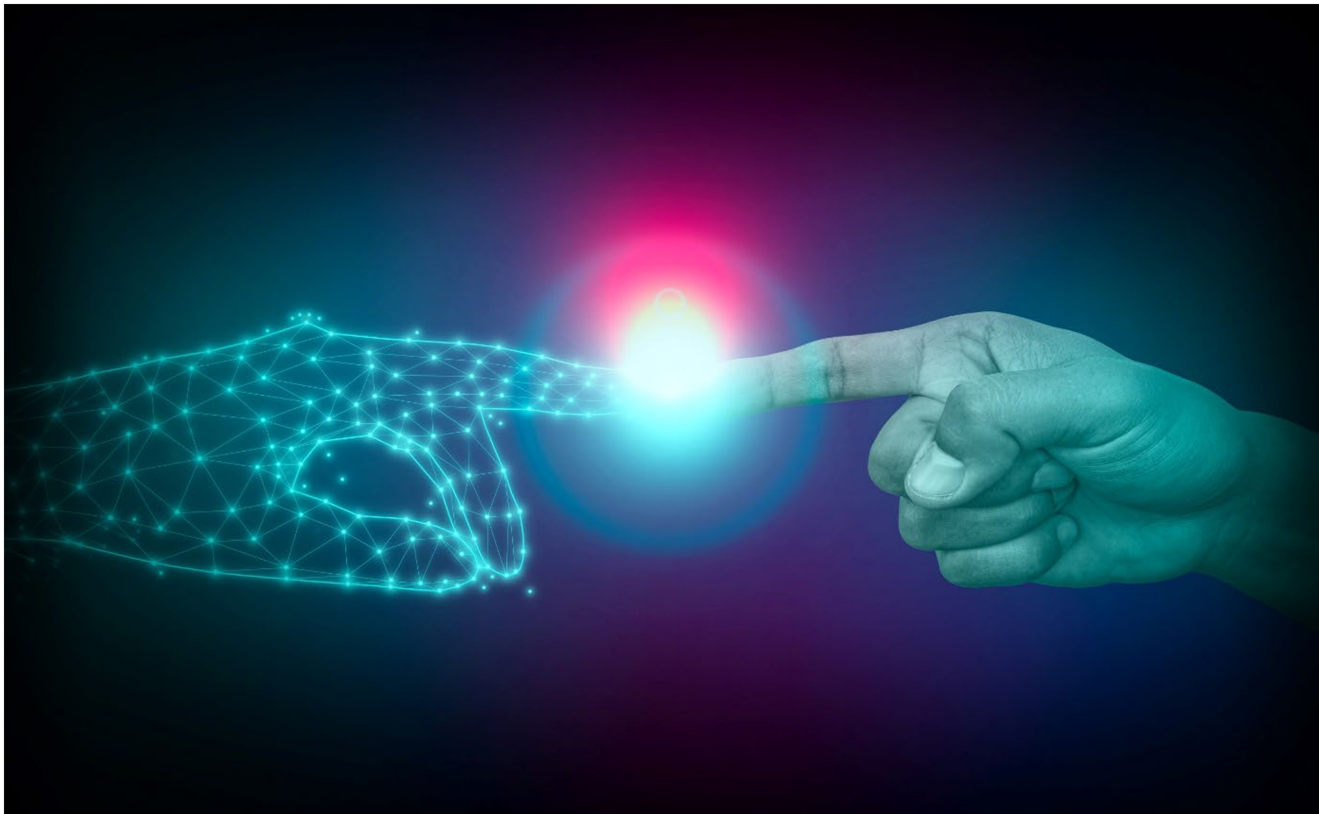
Machine Learning and AI for Healthcare



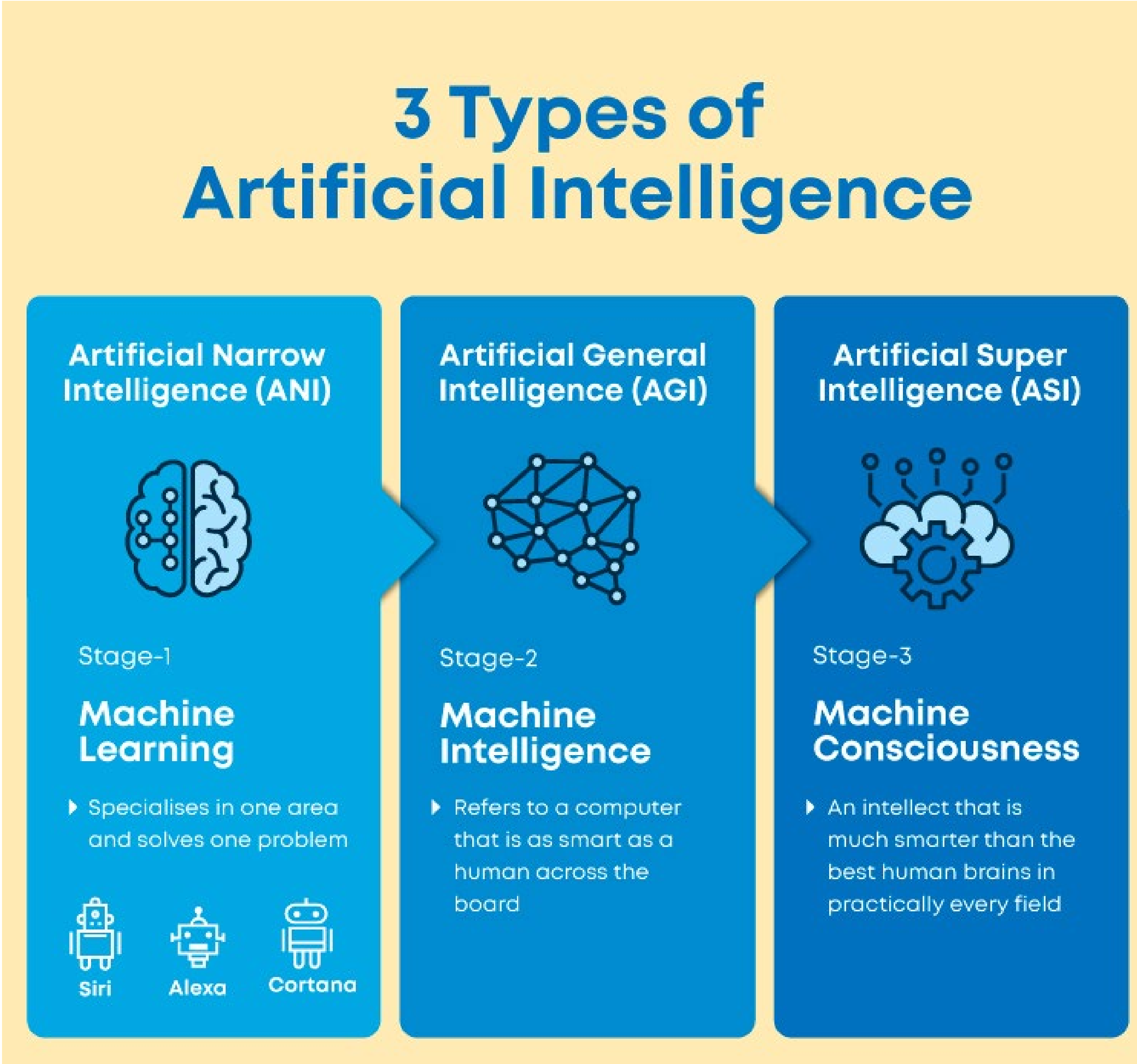
David J Patrishkoff

Feb 11, 2022

7 Presentation Topics

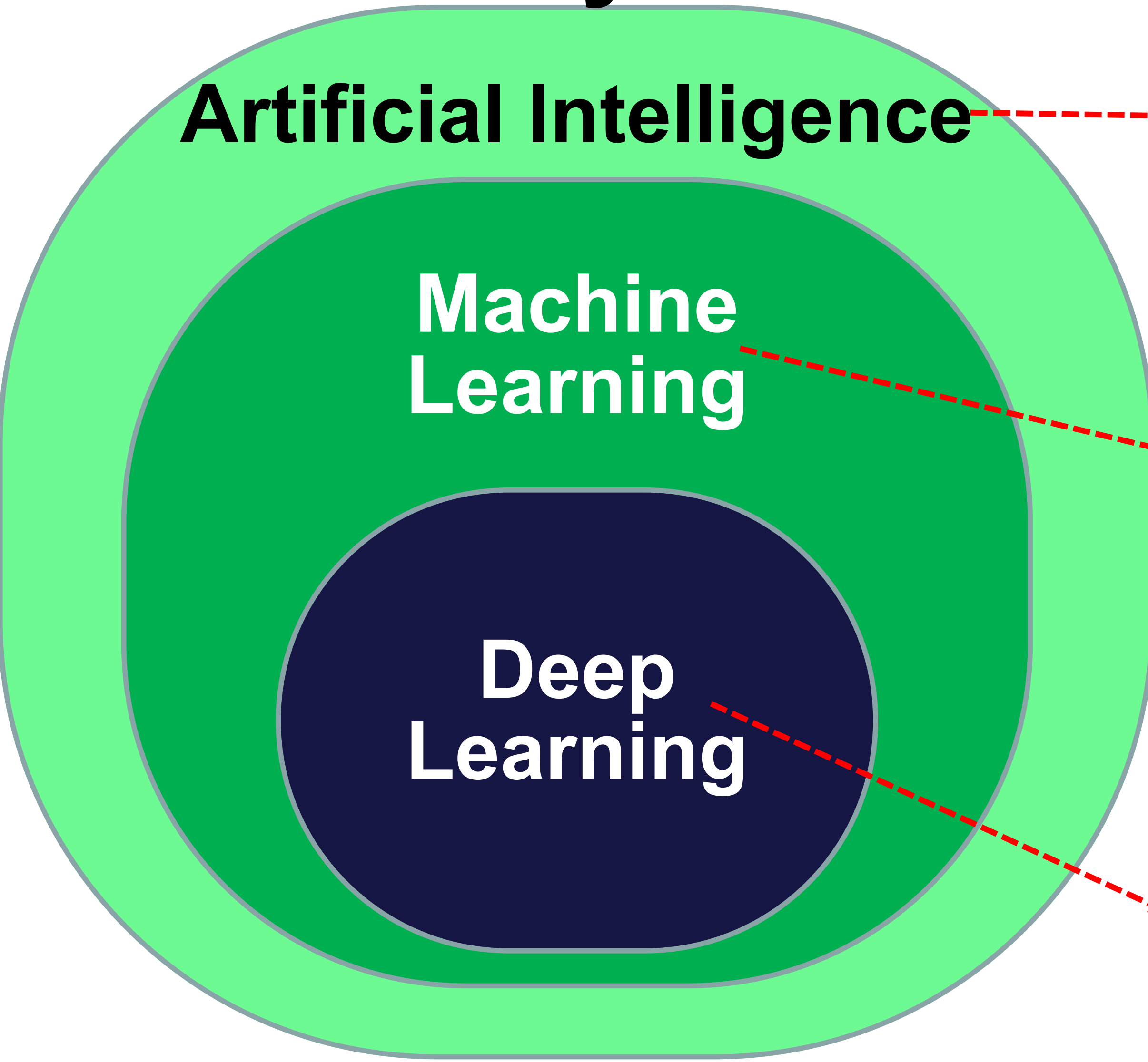
1. What is ML/AI
2. The ML/AI Skills Gap
3. The Home Healthcare Challenge
4. Text Analysis
5. Sentiment Analysis
6. Text Regression Analysis

1. What is ML / AI



https://www.mygreatlearning.com/blog/what-is-artificial-intelligence/

Why AI?... *because*



We all need an ultra-smart personal assistant to improve our quality of life

Smart Algorithms can improve root cause analysis, identify hidden groupings (Clusters) in data and text, and make future forecasts

Computers and their connected devices can constantly learn, re-train themselves and keep improving to meet our needs

Machine Learning is the Engine of AI

Public Sentiment Concerning AI

When a Pega Global AI customer survey asked 6,000 adults in North America, Europe, Middle East, and Africa, and the Asia-Pacific region about AI, confusion and trust issues were exposed.

Respondents were **asked if they thought they have ever interacted with some sort of AI**.

- Only 33% said “yes” even though follow-up questions showed the researchers that 77% of the customers had already unknowingly interacted with AI.
- The same survey noted that only **35% of customers were comfortable** with a business using AI to interact with them and **28% were uncomfortable** with interacting with AI.
- The remaining **37% were undecided**. https://www.pega.com/ai-survey
- A 2021 Hyland survey of 1,000 consumers showed that 57% of consumers believe that “AI has the potential to do damage over the next 10 years due to misuse”. https://blog.hyland.com/hyland-news/digital-distrust-poll-a-healthier-society-starts-with-trust-in-technology/

Algorithmic Discrimination and Bias in Healthcare

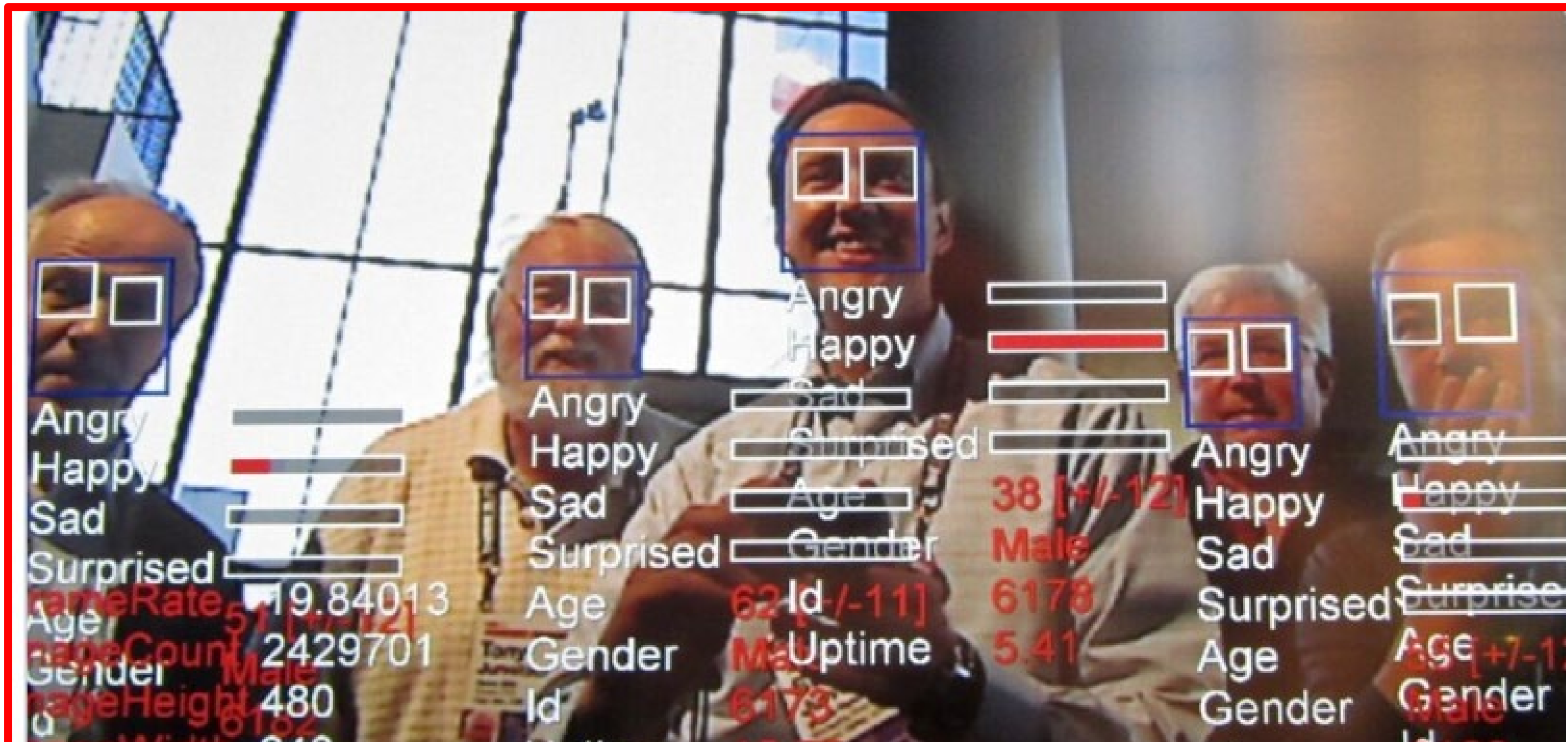
- Physicians use many algorithms that assign **risk scores** for patients.
- Those **risk scores** determine when certain levels of care are taken as well as who gets a organ transplant.
- The **algorithms are adjusted and “corrected**” based on a patient’s race and ethnicity. Such **race-adjusted algorithms** guide the decisions that may direct more attention or resources to white patients than to racial and ethnic minorities. https://www.nejm.org/doi/pdf/10.1056/NEJMms2004740
- In many such cases, “**black” means lower risk** which can direct medical care away from those patients.
- The developers of these algorithms offer no explanations for why these risk score adjustments exist for specific racial and ethnic minorities.

Algorithmic Bias in Healthcare

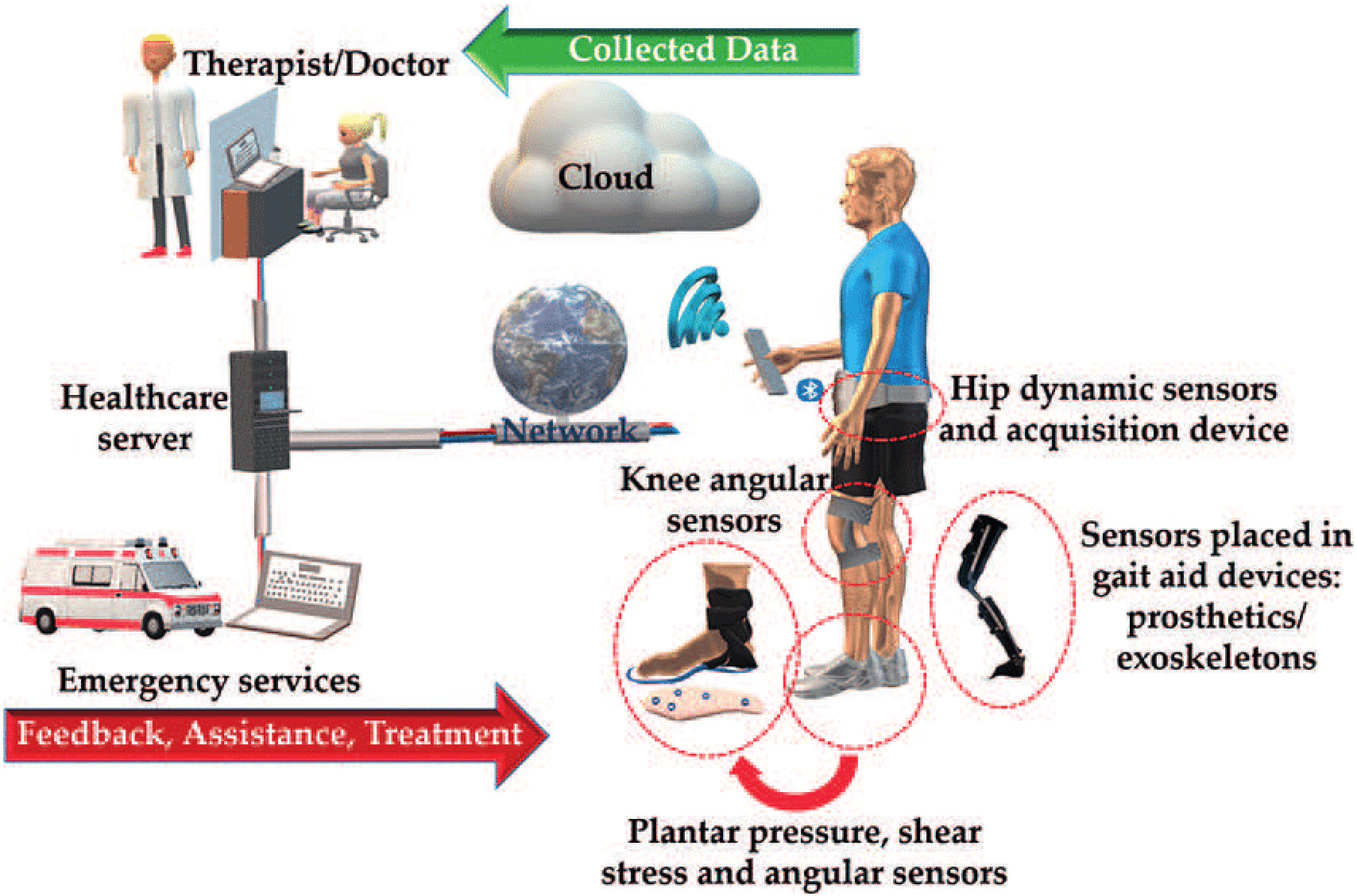
- **2,606 studies** that were published and peer-reviewed in **PubMed between Jan 1, 2015, and Dec 31, 2019**, trained a deep learning algorithm to **perform an image-based diagnostic tasks across 6 clinical disciplines.**
- As broad as this scope may seem, the US patient data came mainly from **California, Massachusetts, and New York** with little or no data from the other **47 states**.
- Data from those few states used to train the algorithm **may generalize poorly to the rest of the nation's patients** due to different economic, educational, social, behavioral, ethnic, and cultural backgrounds. https://jamanetwork.com/journals/jama/fullarticle/2770833?guestAccessKey=ad8f72ad-8b98-42fa-87c3-58c7112d923f&utm_source=For_The_Media&utm_medium=referral&utm_campaign=ftm_links&utm_content=tfl&utm_term=092220

Artificial Emotional Intelligence

- Artificial Emotional Intelligence or Emotion AI is a branch of AI that allow computers to understand human non-verbal cues such as body language and facial expressions.
- This can also be applied to patients
- Affectiva offers cutting edge emotion AI tech: https://www.affectiva.com/



AI-aided Gait Analysis



XRAY Diagnoses

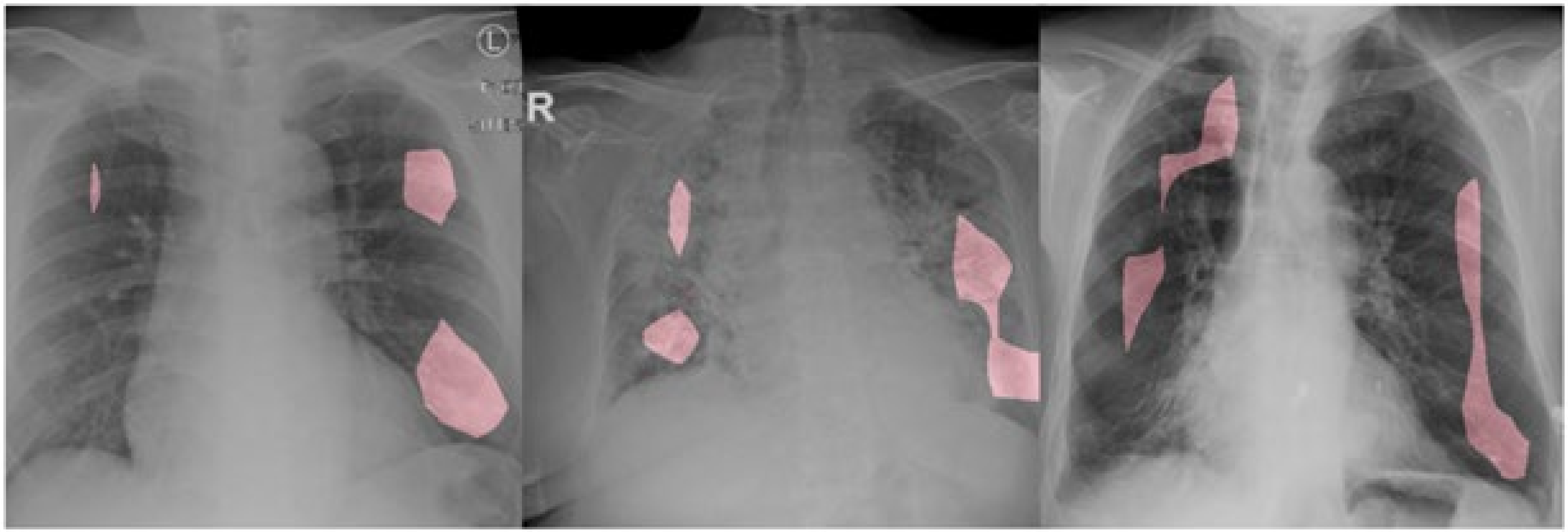


Figure 7. Example CXR images of COVID-19 cases from several different patients and their associated critical factors (highlighted in red) as identified by GSInquire[37].

AI facial recognition capabilities

AI powered facial recognition algorithms have already shown that they can more accurately detect the following human traits with just one picture better than general human assessments as noted below:

- Political orientation internationally – **conservative versus liberal** (73% AI vs 63% Human assessments)
- **Sexual orientation** (76% AI vs 56% Human) assessments
- **Personality type** (64% AI vs 57% Human assessments)
- Other studies are attempting to use facial recognition to identify the level of **honesty**, **intelligence**, and the **tendency towards violence** in individuals, just from a facial picture. https://www.nature.com/articles/s41598-020-79310-1 Although such studies document impressive findings, the risks associated with this technology is that if any decisions and judgments made with this tech about other people will have accuracy levels that are still far from being acceptable or perfect.
- Studies have shown that algorithms trained with 35,326 facial images and **5 images** per person can have accuracy levels of **91% (for men) an 83% (for women**) in identifying sexual orientation. DOI 10.17605/OSF.IO/ZN79K https://osf.io/zn79k/

Quotes about AI

Vladimir Putin (2017):

"Artificial intelligence is the future, not only for Russia, but for all humankind, It comes with colossal opportunities, but also threats that are difficult to predict. Whoever becomes the leader in this sphere will become the ruler of the world."

https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world

Quotes about AI

Stephen Hawking (2018):

“The development of full artificial intelligence could spell the end of the human race….It would take off on its own, and re-design itself at an ever-increasing rate. Humans, who are limited by slow biological evolution, couldn’t compete, and would be superseded.”

https://economictimes.indiatimes.com/news/science/stephen-hawking-warned-artificial-intelligence-could-end-human-race/articleshow/63297552.cms?from=mdr

Quotes about AI

Elon Musk (2014):

"I think we should be very careful about artificial intelligence. If I had to guess at what our biggest existential threat is, it's probably that. So we need to be very careful," said Musk. "I'm increasingly inclined to think that there should be some regulatory oversight, maybe at the national and international level, just to make sure that we don't do something very foolish. With artificial intelligence we are summoning the demon. In all those stories where there's the guy with the pentagram and the holy water, it's like yeah he's sure he can control the demon. Didn't work out."

https://bigthink.com/technology-innovation/elon-musk-we-should-be-very-careful-about-artificial-intelligence/

Quotes about AI

Masayoshi Son (2017), chief executive of SoftBank Group Corp, said:

"Those who rule data will rule the world".
https://www.fintechfutures.com/2017/07/softbank-gets-hard-for-data-powered-ai/

Quotes about AI

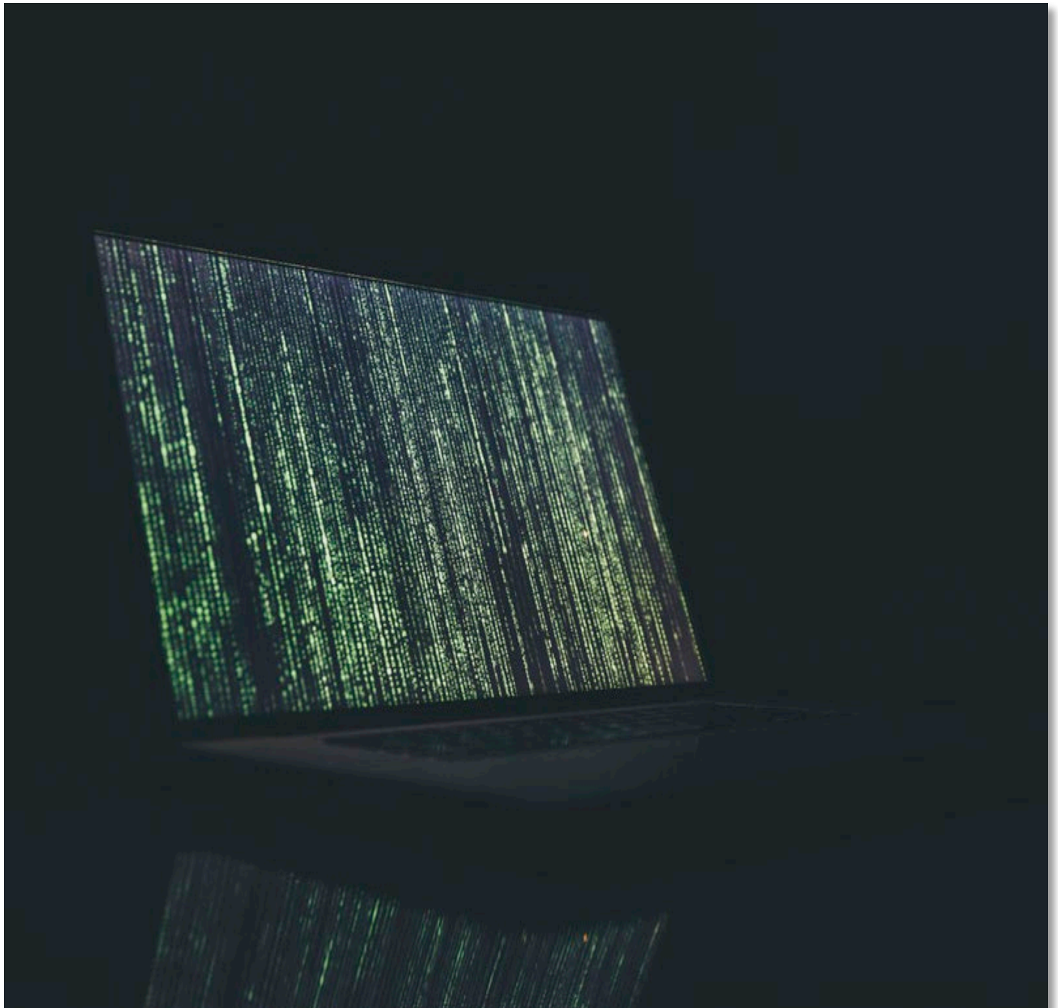
The Brookings Institute has said:

“Whoever leads in artificial intelligence in 2030 will rule the world until 2100.”

https://www.brookings.edu/blog/future-development/2020/01/17/whoever-leads-in-artificial-intelligence-in-2030-will-rule-the-world-until-2100/

2. The ML / AI Skills gap

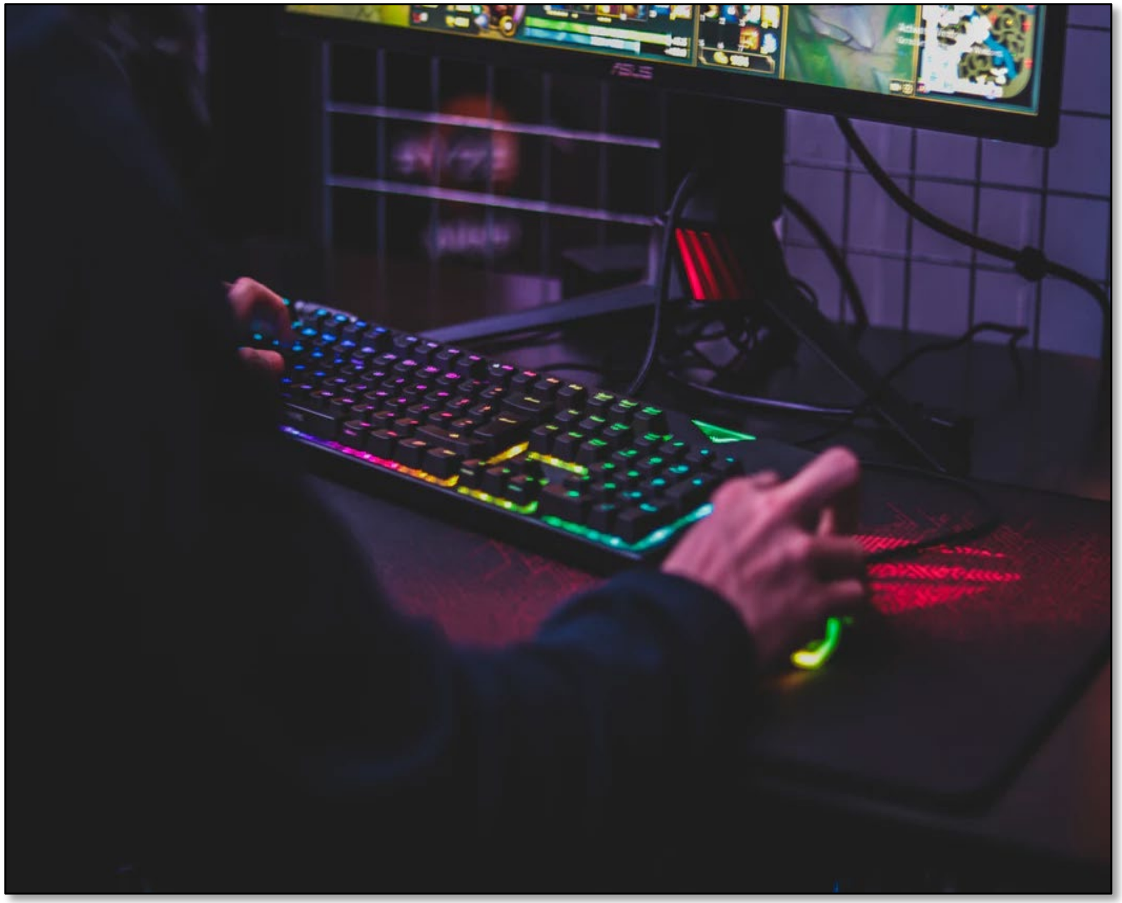
The Shortage of Data Scientists



- 2018 LinkedIn Workforce Report found that there were more than 151,000 data scientist jobs unfilled across the US

- According to the recent KPMG CIO Survey, the biggest skill shortages are for:
    - Big Data analytics
    - AI skills

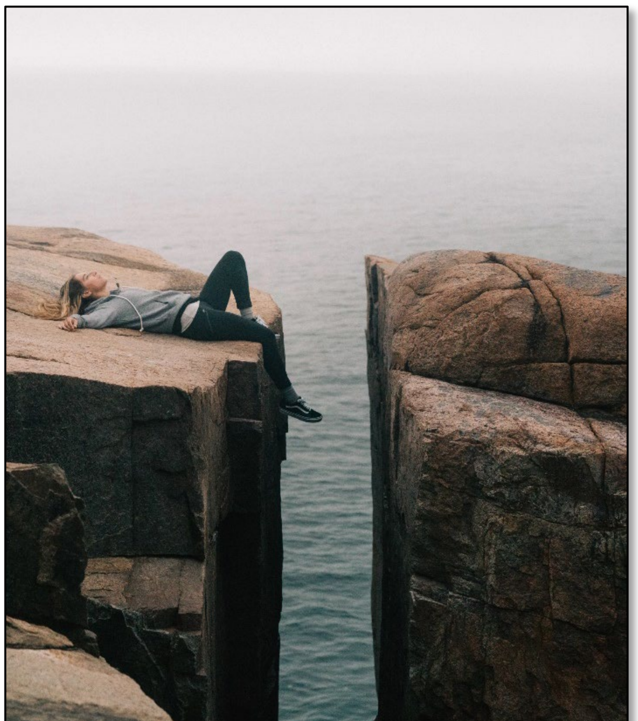
Advantages of Adding ML Skills to your Personal Toolkit



- Acquiring Machine Learning skills is like adding an insurance policy to your career
- ML skills can protect you against certain disruptive changes
- It is safer to be on the side of being the potential *Disrupters* compared to being with the *Disrupted*

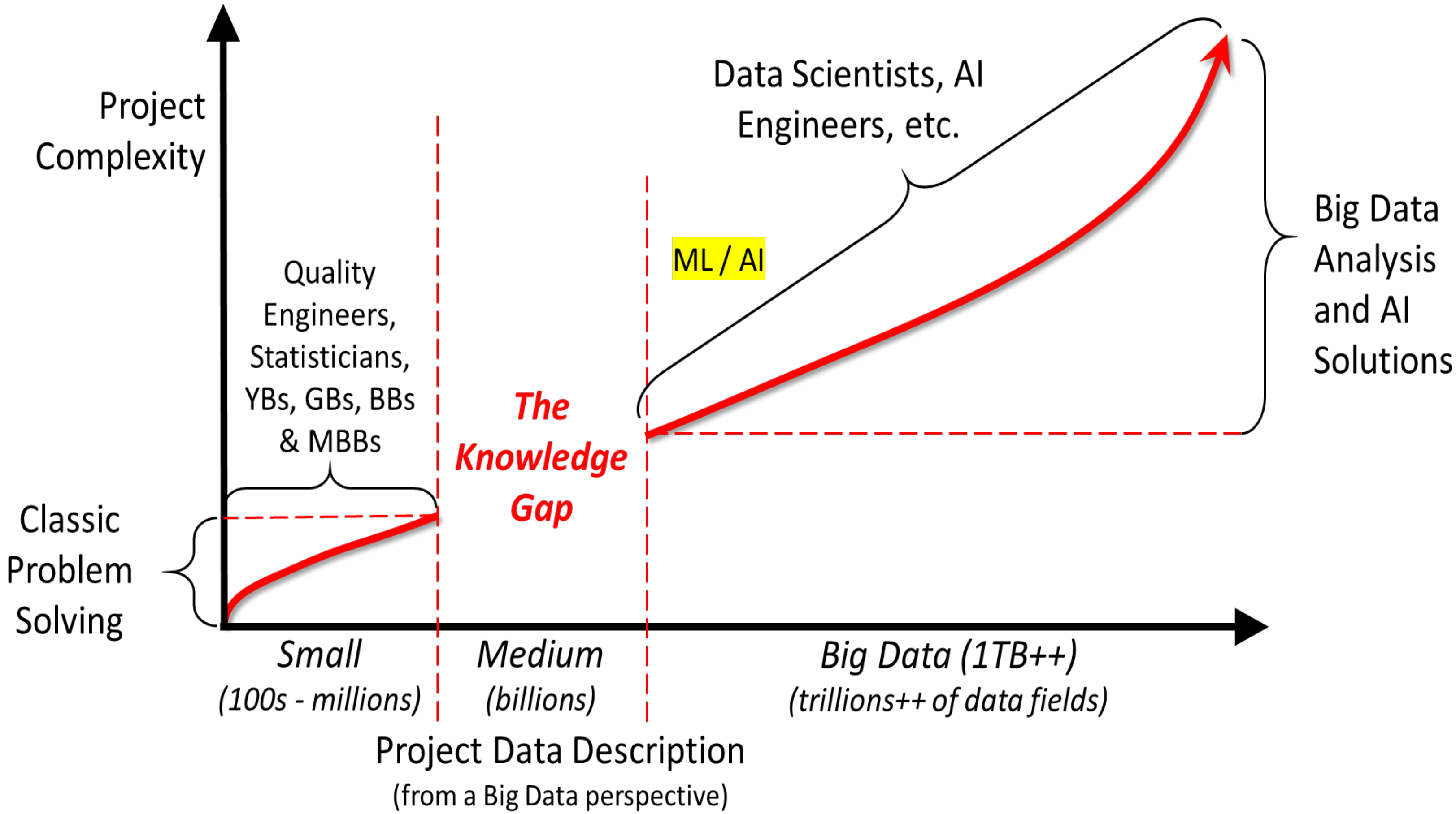
How Wide is this Knowledge Gap Between Statistics & ML/AI?

Narrow?? < OR > Very Wide??



Robbie Maddison jumping the Corinth Canal in Greece (2010)

The Knowledge Gap Between Basic Statistics & ML



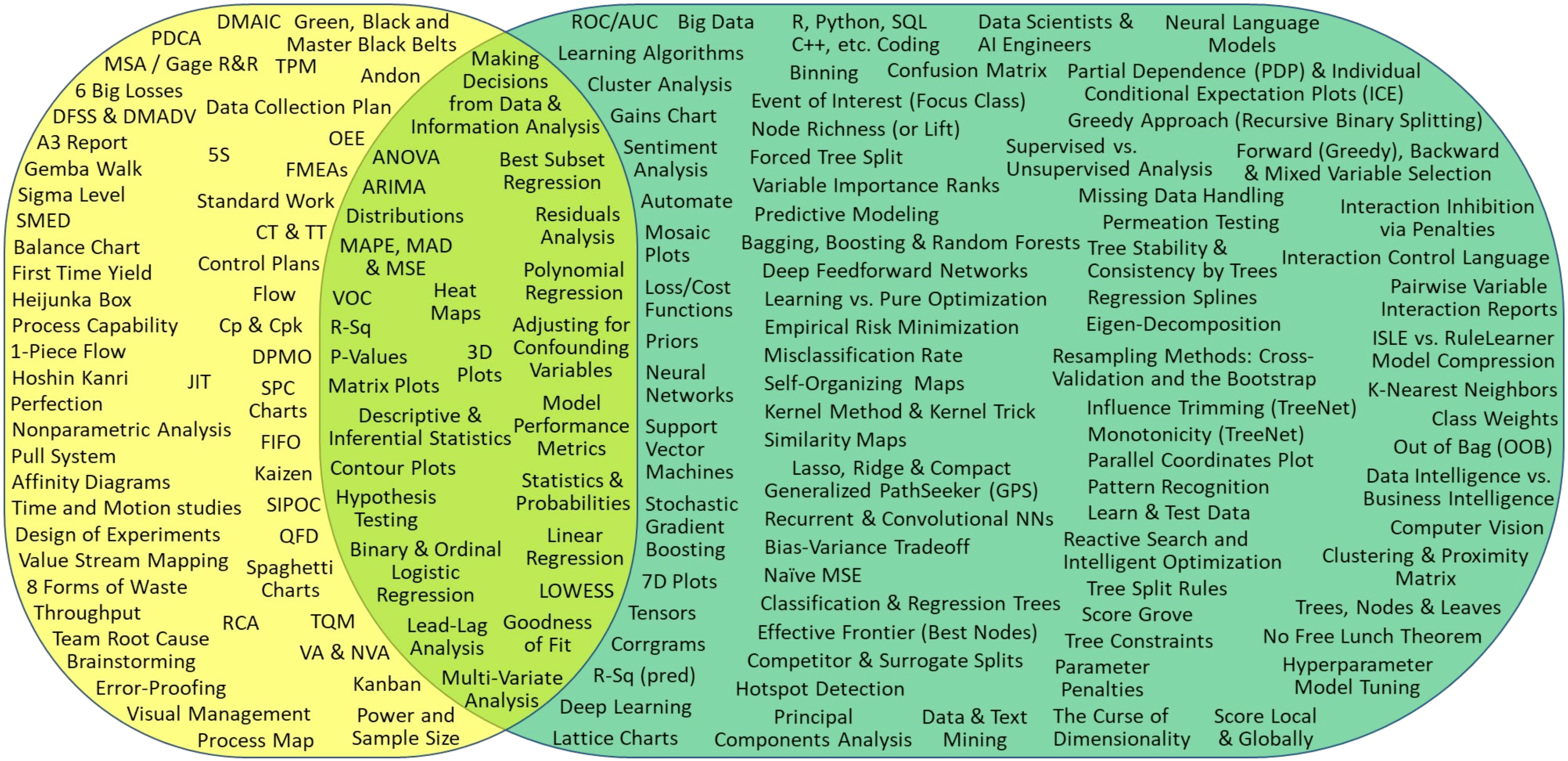
Statisticians and Six Sigma practitioners need new tools to address the ever-increasing complexity of business data

“Citizen Data Scientists” can Bridge the Knowledge Gap Between Staistics & Data Science

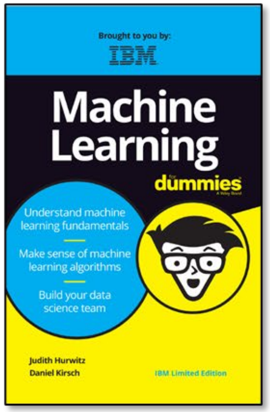
- A Data Scientist position is filled 90% of the times with someone who has either a Master’s degree or PhD in Data Science
- In 2016, Gartner coined the phrase: Citizen Data Scientist (CDS) to reflect those individuals who could assist in the data analysis revolution with advanced software tools that do not require coding
- CDS positions offer Statisticians additional career engagement paths

Unique & Shared Vocabulary Between the LSS/Basic Statistics & ML Worlds

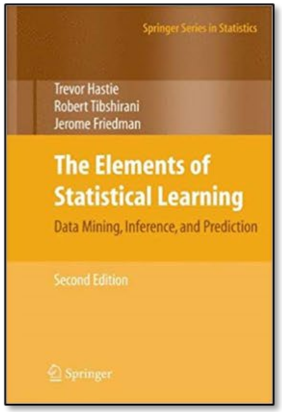
Some Stats Vocabulary **Some ML Vocabulary, excluding text analysis**



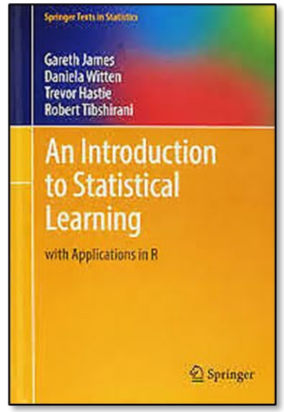
9 of the Many Great Machine Learning Books



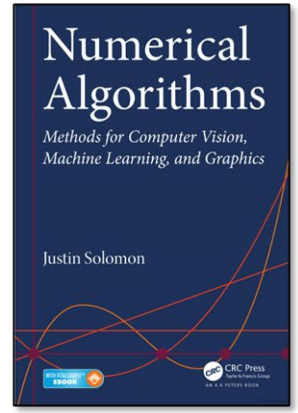
IBM
75 pages



Springer Press
764 pages

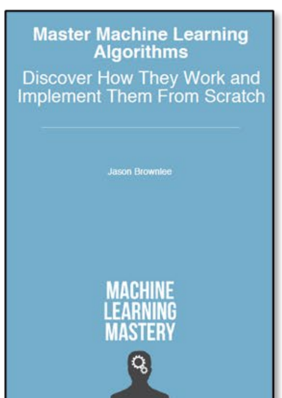


Springer Press
441 pages

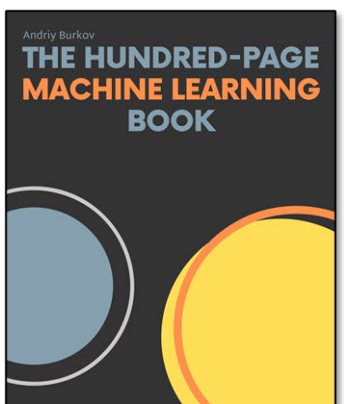


CRC Press
397 pages

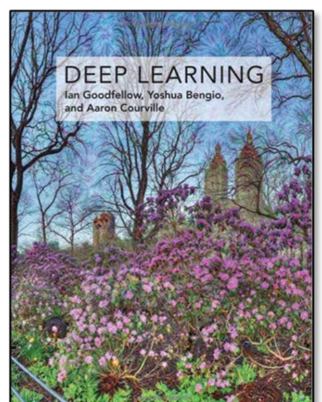
Packt Press
272 pages



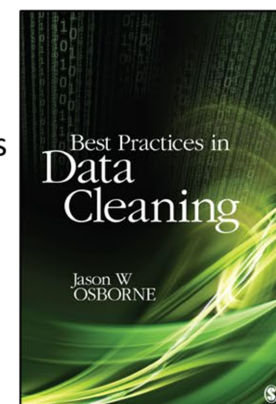
< Jason Brownlee
163 pages



< Andriy Burkov
100 pages

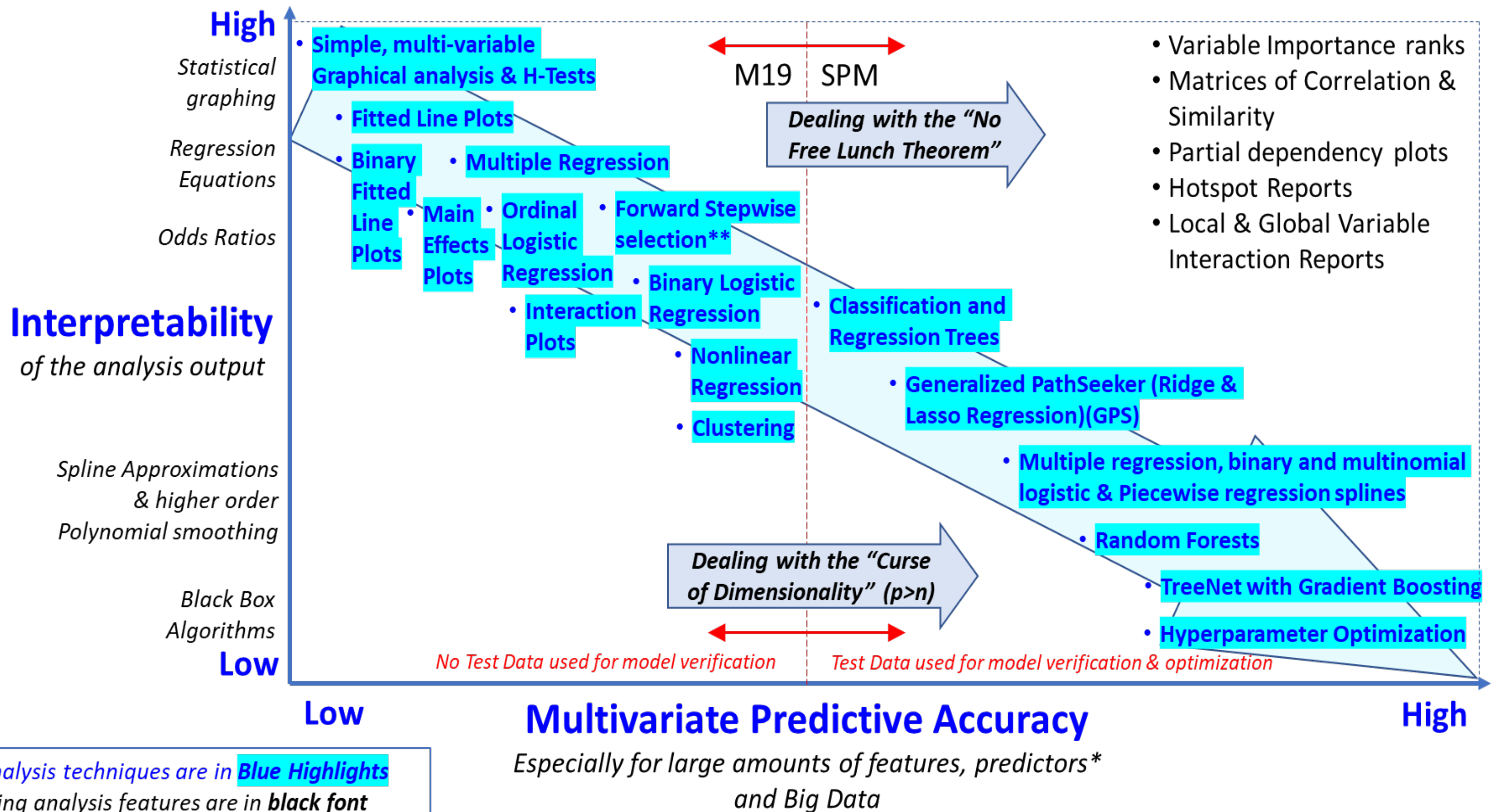


< MIT Press
850 pages



< Sage Press
293 pages

Various Methods for Standard Stats Analysis vs. ML Data Analysis

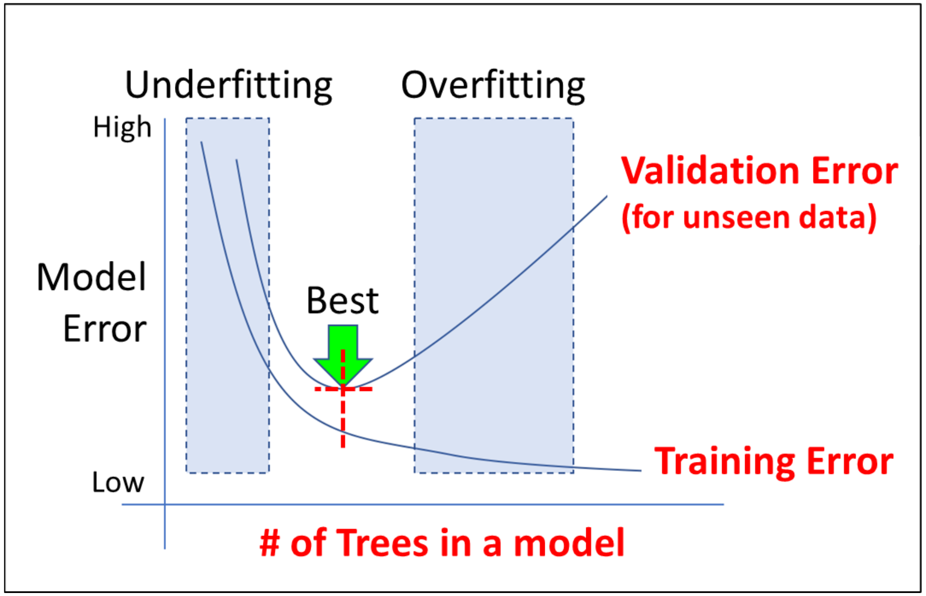


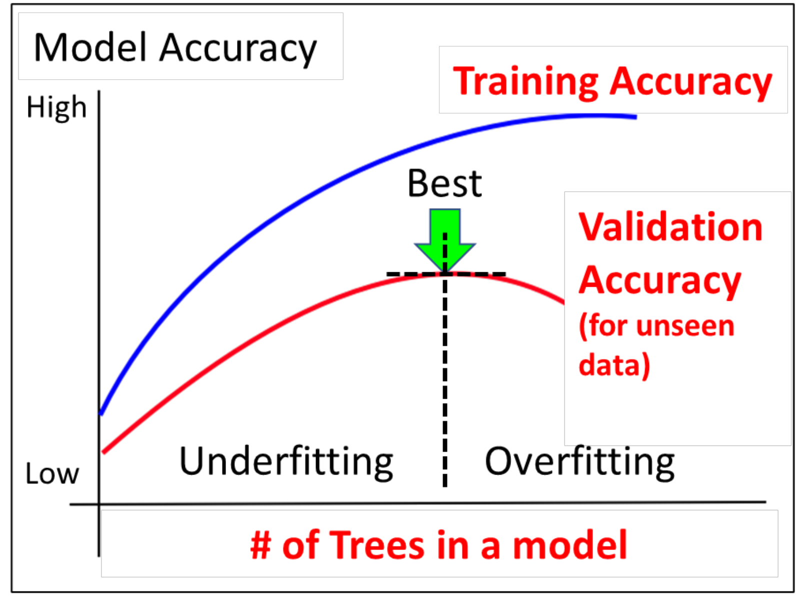
Main Analysis techniques are in ***Blue Highlights***
Supporting analysis features are in ***black font***

- James, G., et.l. (2013), If *n (# of observations)* is not much larger than *p* (# of predictors), then there can be a lot of variability in the least squares fit, resulting in overfitting and consequently poor predictions on future observations not used in model training. Page 204

** Unlike Best Subset reg that is limited to 40 p max, and backward stepwise cannot deal with p>n, forward stepwise can be used even when $n < p$, and so is the only viable subset method when p is very large. Pg 208

Accuracy of a Model – Training vs. Validation Data





ML software will automatically calculate and report the conditions required to achieve the best **Validation** performance metrics and lowest errors for a model for unseen data

3. The Home Healthcare Challenge

Home Healthcare Research Scope & Goals

Scope: All patients in the USA that are being cared for by nurses and doctors from standalone Home Healthcare Agencies, not hospitals or assisted care facilities.

Goals: Support "Aging in Place" goals by dramatically improving Home Healthcare performances.

Provide the best intervention recommendations to Nurses in Home Healthcare agencies.

Apply Predictive Modeling techniques to validate and predict best practices hidden in millions of past intervention data for very specific sub-sets of patients.

What is Aging in Place?

Staying in your own home as you get older is called “Aging in Place“ and it the preferred scenario by the elderly. Typical Comments from the elderly:

- I've lived here 40 years. No other place will seem like home.
- The stairs are getting so hard to climb.
- Since my wife died, I just open a can of soup for dinner.

National Institute on Aging

The Elderly Quality of Life Contradiction Statement

"I really want to keep living in my home…...

but I might not be do that safely as my health and level of independence declines with increased age"

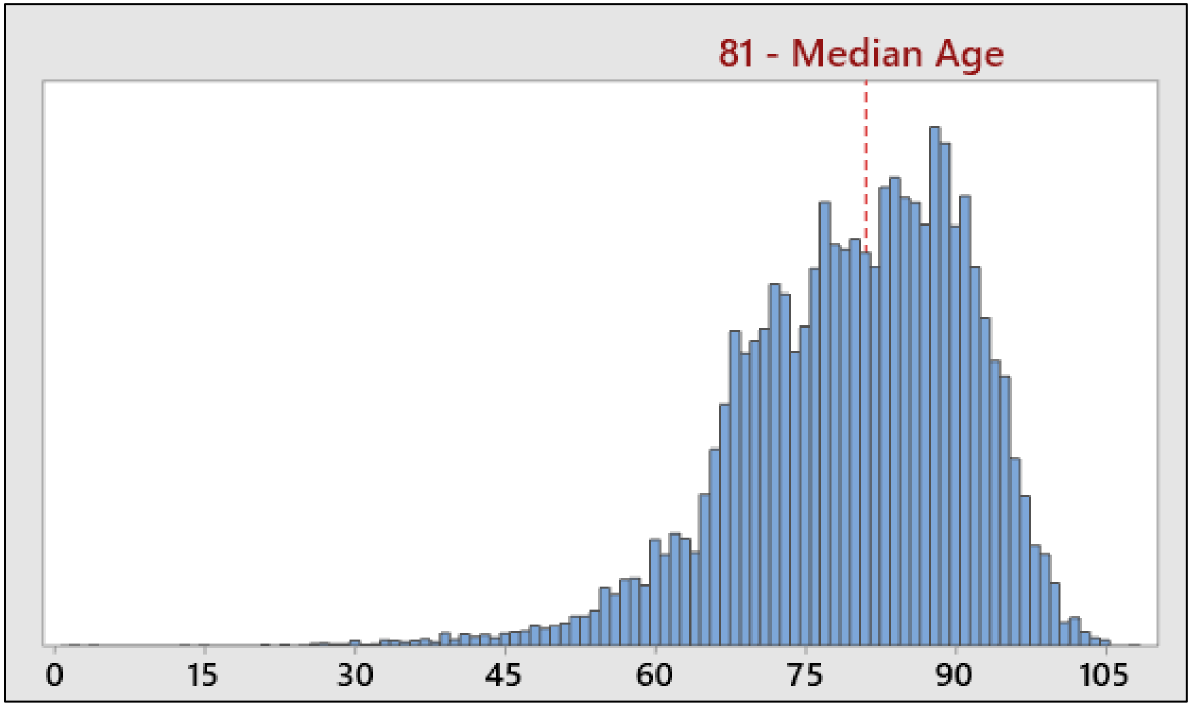
Home Healthcare Supports "Aging in Place" Goals

Home Healthcare is a wide range of health care services that can be given in your home for any level of illness or injury.

It is usually less expensive, more convenient, and just as effective as care you get in a hospital or skilled nursing facility.

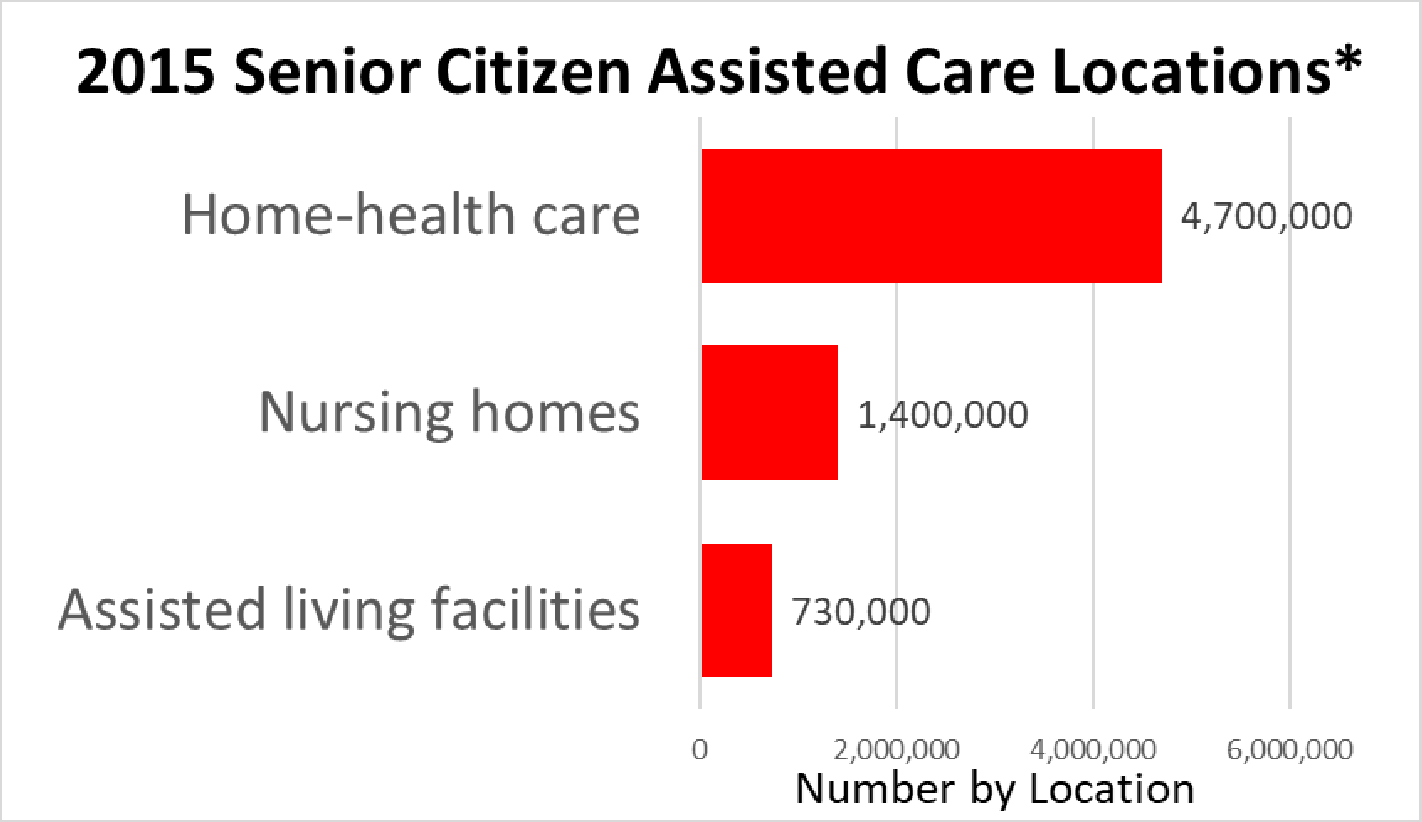
It is only viable if patients can safely perform specific Activities of Daily Life (ADLs) at home without the support of fulltime skilled nursing assistance.

Age of Home HC Patients



- Home Healthcare assists patients of all ages and for all severity levels of illnesses and disabilities.
- Median Age of a Home Healthcare patients is 81.
- When Home Healthcare efforts fail, hospitals and nursing homes are typically the next option.

Senior Citizen Care Locations



Home Health care is the last line of defense before other expensive and inconvenient health care options and locations must be considered

*National Institute on Aging

Activities of Daily Life (ADLs)

- Grooming
- Dress Upper & Lower Body
- Bathing
- Toilet Transferring independently
- Ability of Transferring to different positions
- Safe Locomotion (Walking or Wheelchair)

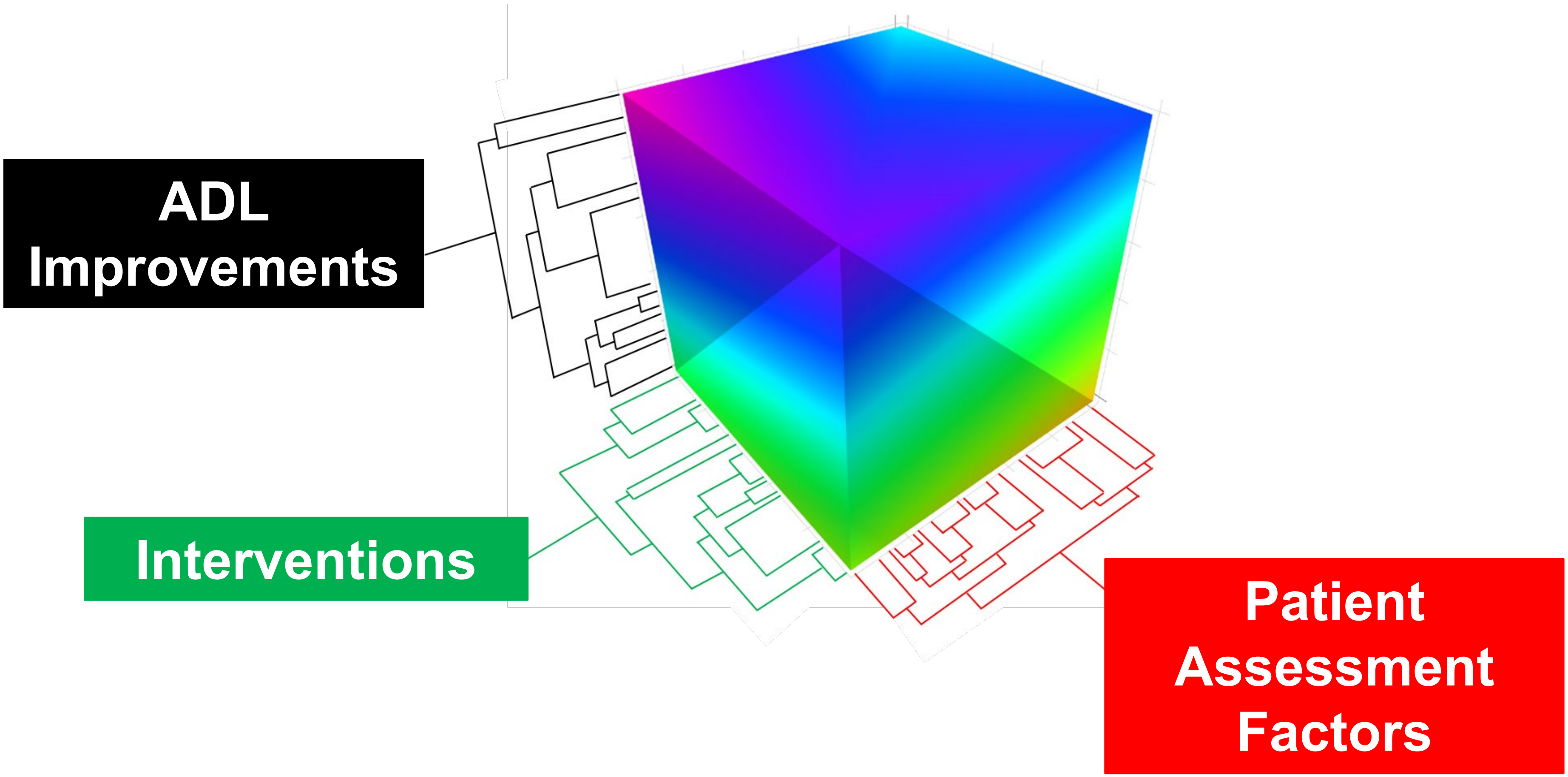
ADLs scores tell us how safe it is to live independently, in your home

How do we know when it is unsafe to "Age in Place"?

- When the combination of a patient's bad ADL Score (determined by a nurse) and the lack of support resources at home signals unsafe independent living conditions

ADLs scores + the level of needed support resources tells us how safe it is for a person to keep living independently in their home

Research Goal: Improving ADL Scores



We are applying Predictive Modeling to assist Home Health Care Nurses to make better Care Decisions

Factors Under Ongoing Analysis

- Before and after Patient ADL Scores
- 83,000+ Medication effects and interactions
- 43,000+ Diagnosis Codes
- 1,200+ Patient Care Interventions
- 100+ Other Patient Factors

The ability to analyze this and other complex healthcare data sets is beyond the reach of classic six sigma statistical tools

Current versus Future State of Home Healthcare

The Current Wild West

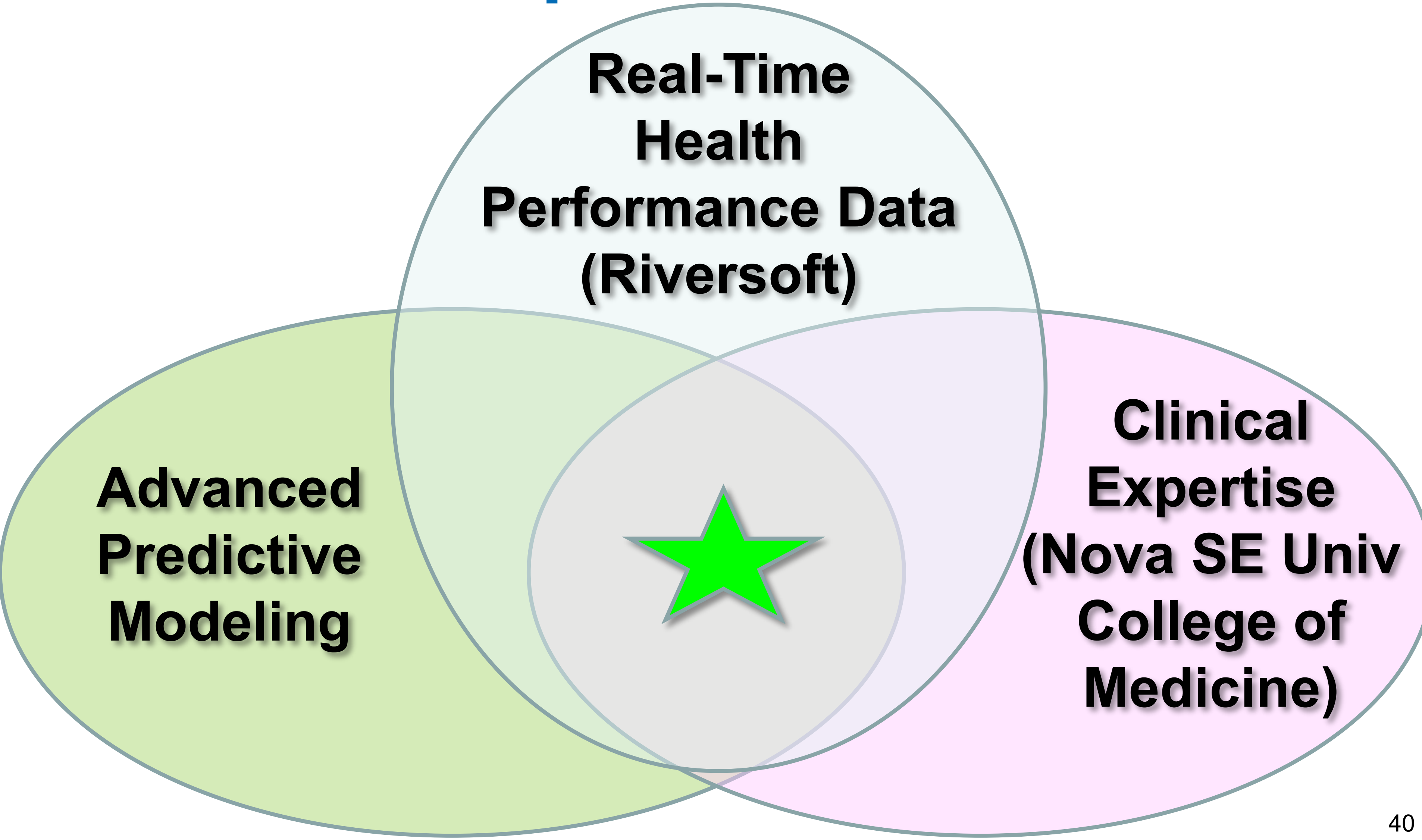
Individual instinct & impulse-driven decision-making

The Future

Self-Improving ML Algorithms based on national data bases

We are currently piloting steps to move to the future

Coordinated Activities to Improve ADLs



Our Source of Data for Home Healthcare Performances

- Since 1997 RiverSoft LLC provides applications and has healthcare data for home health and hospice agencies in over 20 states
- Data analysis and Predictive analytics were applied to this data set to determine the best care interventions that achieved the best ADL outcomes for patients
- This data is used for billing and scrutinized by Medicare for accuracy

The above-noted data is the largest central set of home healthcare performance data in the USA

Analysis of 1,259 Current Home HC Interventions

- **157** have *always* attained an improvement of 5 or more on the ADL Scale (Best of the Best)
- **90** have *always* achieved an improvement between +1 and +4 (Good Results)
- **74** have *always* attained a Zero improvement
- **14** have *always* had a negative outcome (Worst of the Worst)
- The remaining **924** interventions had very mixed results

We are alerting Home Healthcare nurses about the best and worst interventions

Pilots: Predictive Modeling Improving the Quality of Life?

Real-Time Home Healthcare Data from RiverSoft covering over 20 states

Predictive Analytics: Determining the best Interventions with the best performance

Data from the Field will show what works and what does not

Improve and Expand the Predictive modeling to improve the Quality of Life

Pilot Launch: RiverSoft ORIB™ with Predictive Modeling

- RiverSoft ORIB™ (Outcome Ranked Intervention Browser™) highlights the best care interventions that influence ADLs based on Predictive Modeling
- Home Healthcare Nurses can now compare their favorite care interventions with a nationwide data base for what actually works best

Expanded Ongoing Predictive Modeling for Home Healthcare

- We are also conducting Predictive Modeling research to reduce:
  - ER visits
  - Fall risk
  - Pain
  - Wound care issues

Our ADL research findings are being applied to the real world and being expanded to other areas

4. Text Analysis

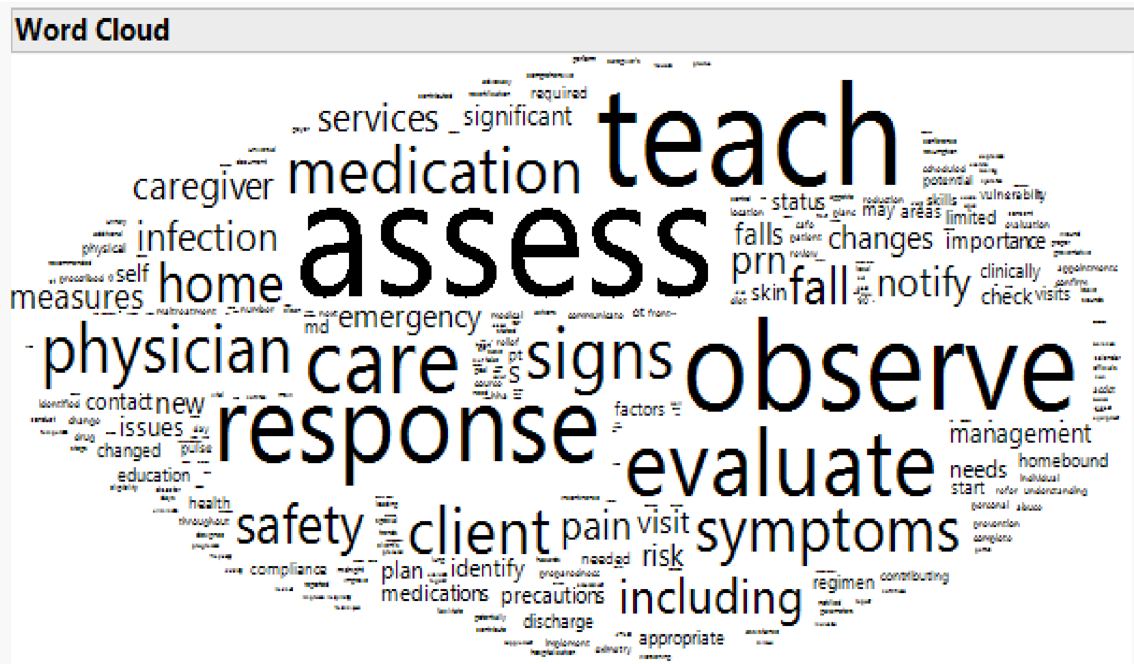
Text Exploration in JMP Pro - Home Healthcare MS rehab data analysis from the largest home Health care agency in this data set- 115,746 interventions

Term and Phrase Lists

Term	Count
assess	49837
teach	41530
observe	36872
response	31497
care	27574
evaluate	27405
signs	21167
physician	20364
client	19446
medication	19394
symptoms	19053
home	16526
safety	16429
fall	15197
including	14894
infection	12384
pain	12375
notify	12233
prn	12123
caregiver	11569
services	11284
measures	11026
changes	9815
emergency	9702
visit	9615

Phrase	Count	N
assess observe	36142	2
evaluate response	23667	2
signs symptoms	12140	2
client caregiver	10978	2
notify physician	9121	2
significant medication issues	8350	3
medication issues	8350	2
significant medication	8350	2
symptoms of infection	6951	3
teach importance	6447	2
clinically significant medication issues	6330	4
new and or changed	6330	4
clinically significant medication	6330	3
changed medications	6330	2
clinically significant	6330	2
signs symptoms of infection	6285	4
start of care	6231	3
including but not limited	6042	4
home care	5563	2
home care services	5307	3
care services	5307	2
signs and symptoms	4788	3
self management	4274	2
areas of vulnerability	4041	3
teach client	4017	2

Common word and term pareto charts and frequency tables



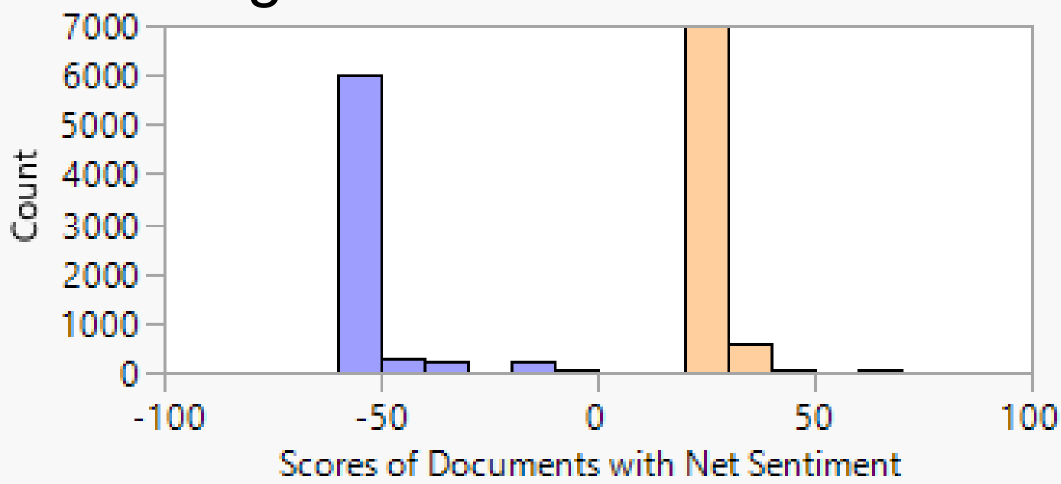
The most common words used in home healthcare interventions are shown with a larger font

5. Sentiment Analysis

Sentiment Analysis Summary in JMP Pro

	N	Mean Score
All Scored Documents	14450	-14.6
Net Positive Documents	7447	22.6
Net Negative Documents	6520	-58.1
No Sentiment Documents	101296	0.0

Negative words Positive words



Sentiment Analysis Summary in JMP Pro

Sentiment	Score ^	Count
:(	-70	3
pain	-60	8022
problem	-60	3
poor	-50	33
upset	-50	7
waste	-50	1
adverse	-40	359
irritation	-40	15
limited	-20	147

Negative words Pareto Chart

All text analysis requires text cleaning, For example, the SW thought that :(was a frowny face expression

```
[25703] ...fever, purulent drainage); utilize Phlebitis Scale for assessment:(0) no phlebitis present; (1) erythema...
[38619] ...fever, purulent drainage); utilize Phlebitis Scale for assessment:(0) no phlebitis present; (1) erythema...
[68369] ...fever, purulent drainage); utilize Phlebitis Scale for assessment:(0) no phlebitis present; (1) erythema...
```

All pareto words can be easily accessed to check it the positive or negative context was correctly interpreted

Sentiment Analysis Summary in JMP Pro

not limited	20	6065
functional	20	863
adequate	25	16
appropriate	30	3281
thrill	40	1
good	60	483
without pain	60	3
essential	60	1

Positive words Pareto Chart

[113486] Teach on maintaining a clean mouth is essential. Brush gently, with a small toothbrush if needed...

6. Text Regression Analysis

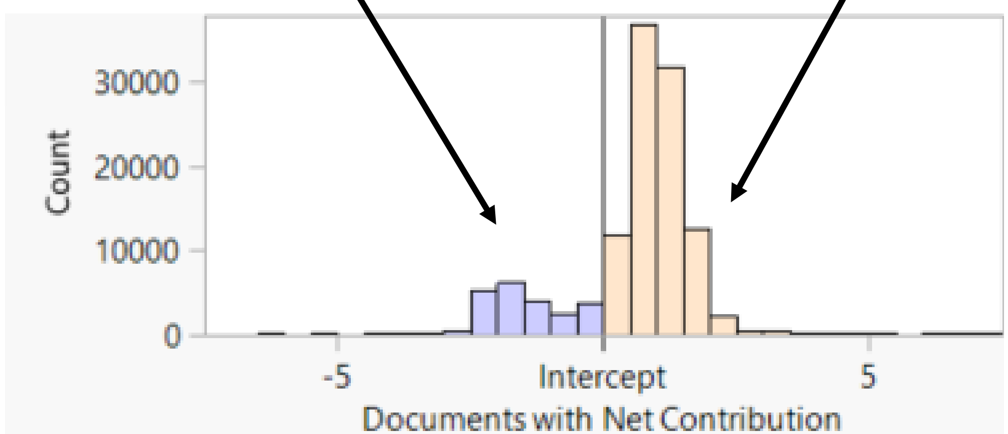
Text Analysis (Term Selection) Generalized Regression for ADL Improvement Summary

	N	Mean Contribution
All Documents	115746	0.568
Net Positive Contribution Documents	94683	5.471
Net Negative Contribution Documents	20975	-21.564
No Contribution Documents	88	0.000

Model Intercept Value: 8.938

Most neg ADL Improvement words

Most pos ADL Improvement words



Text Analysis (Term Selection) Generalized Regression for ADL Improvement Summary

Pareto chart of the most positive ADL Improvement words

Term	Coefficient	LogWorth	Count
pharmacological	2.563	0.965	1012
oral	2.692	0.956	494
evaluate	2.700	51.191	27405
implement	2.714	0.760	3292
ideation	2.921	1.324	878
evaluation	2.931	5.720	3612
elimination	2.995	1.166	2038
lower	3.076	0.546	880
day	4.042	5.644	3329
prescribed	4.088	3.877	3309
evalulate	4.612	0.820	1641

```
[231] Teach importance of balanced diet and/or nutritional supplement and evalulate response.
```

All pareto words can be easily accessed to check it the positive or negative context was correctly interpreted

Text Analysis (Term Selection) Generalized Regression for ADL Improvement Summary

Pareto chart of the most negative ADL Improvement words

Term	Coefficient	LogWorth	Count
dosing	-6.387	14.775	584
iadl	-4.059	0.281	1414
duration	-3.939	1.083	704
upgrade	-3.131	3.962	745
improve	-3.035	1.606	3051
relief	-2.987	8.417	3529
teaching	-2.954	78.129	1585
missed	-2.953	1.931	2187
green	-2.891	0.903	488
msw	-2.705	1.822	3056
extremities	-2.700	0.460	877

[1428] Check/draw INR on 3/15/19, Current dosing 15 mg on Tues 3/12/19 then 10 mg...

[1669] ...diet, home glucose testing, oral and/or insulin dosing and exercise.

[1671] ...exercise, home glucose testing, oral and/or insulin dosing, diabetic precautions, sharps disposal, complications of non...

[1706] ...diet, home glucose testing, oral and/or insulin dosing and exercise.

[1707] ...exercise, home glucose testing, oral and/or insulin dosing, diabetic precautions, sharps disposal, complications of non...

Q & A

Any Questions?

Author contact information:

David.Patrishkoff@CascadeEffects.com