Consideration of Artificial Intelligence Application and Impact for the Electric Power Gas Turbine Industry

ASME Turbo Expo GT2019-91963

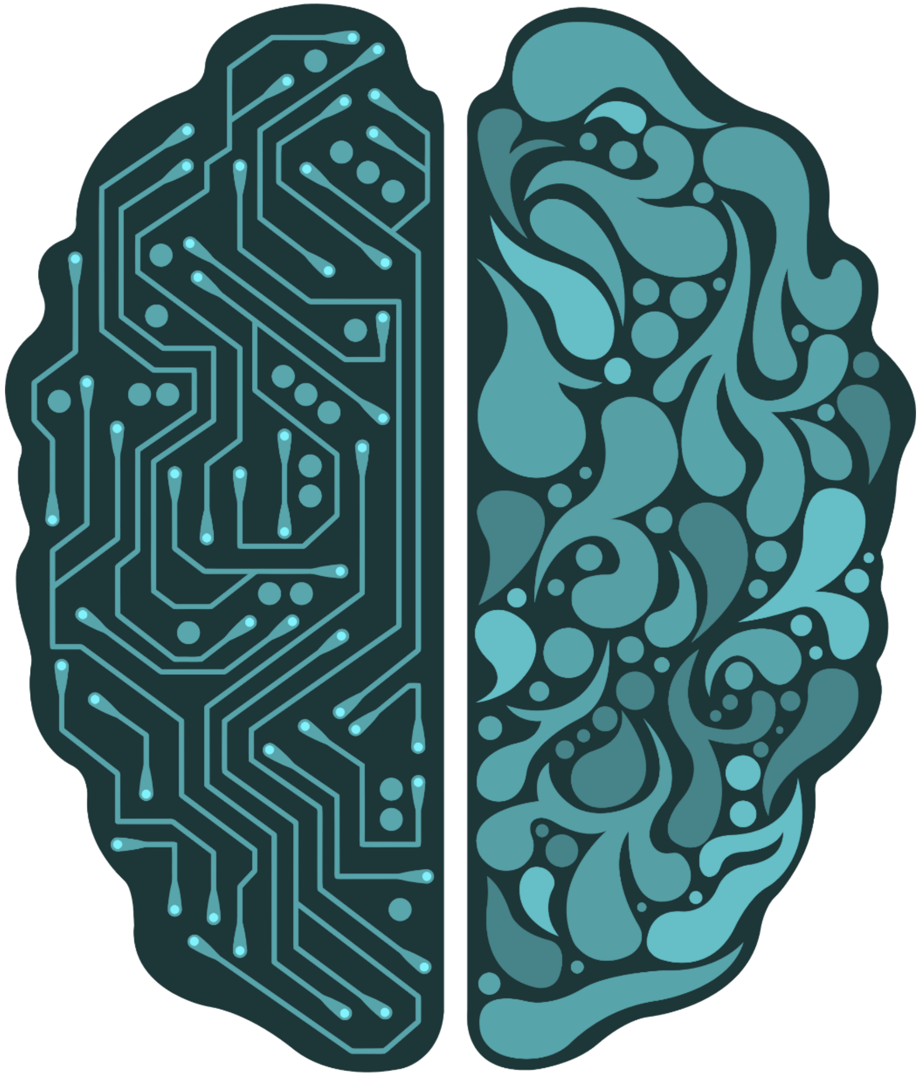
Bobby Noble, Steve Seachman, Leonard Angello, Neva Espinoza
Electric Power Research Institute

Chris Perullo, Ben Emerson, Tim Lieuwen
Georgia Tech

Phoenix, Arizona
June 20, 2019

www.epri.com

Artificial Intelligence



Whatever Computers Can't Do

…Until They Can

By the Numbers

3096 **AI companies with 55 major investors**

2 **Billion Dollars invested in ML startups between 2006 and 2016**

1100 **New AI startups have raised their first round of equity funding since 2016**

505 **Million Dollars invested by Corporate Venture Capital in 2016**
An 80% growth as compared to 2015

But…Where do we go from here?

“You’ve got to be very careful if you don’t know where you are going, because you might not get there.”

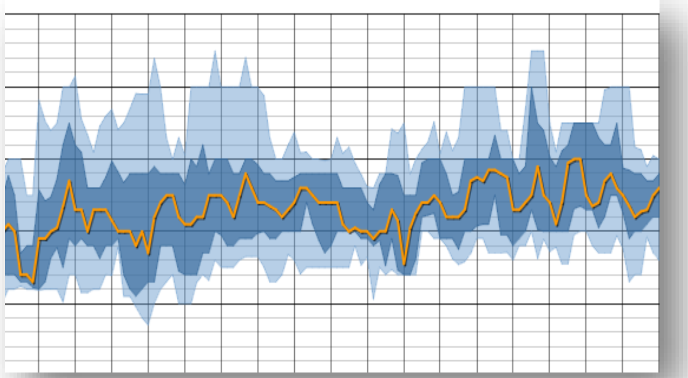
Yogi Berra
The Yankees

Information through Analysis: Spectrum of Data Analytics

Operational

Computer-Assisted SME

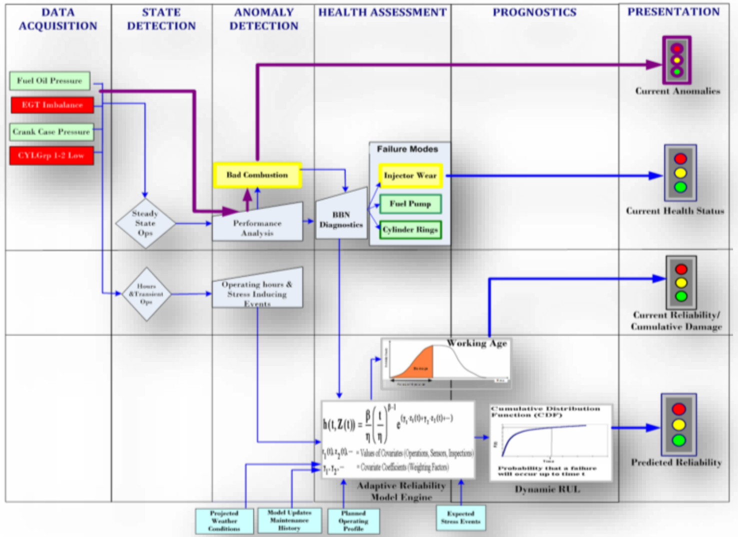
- PI
- Advanced Pattern Recognition (APR)
- Model Development
- Alarm Limits



Predictive

SME-Validated

- Advanced Infrastructure
- Edge Analytics
- Prognostics / RUL



Prescriptive

Computer-Guided SME

- Integrated Technical / Environment / Business Data
- Embedded Complexity

Gas Turbine: Typical Monitoring & Challenges

Instrumentation Problems

Tuning/Performance Deviations

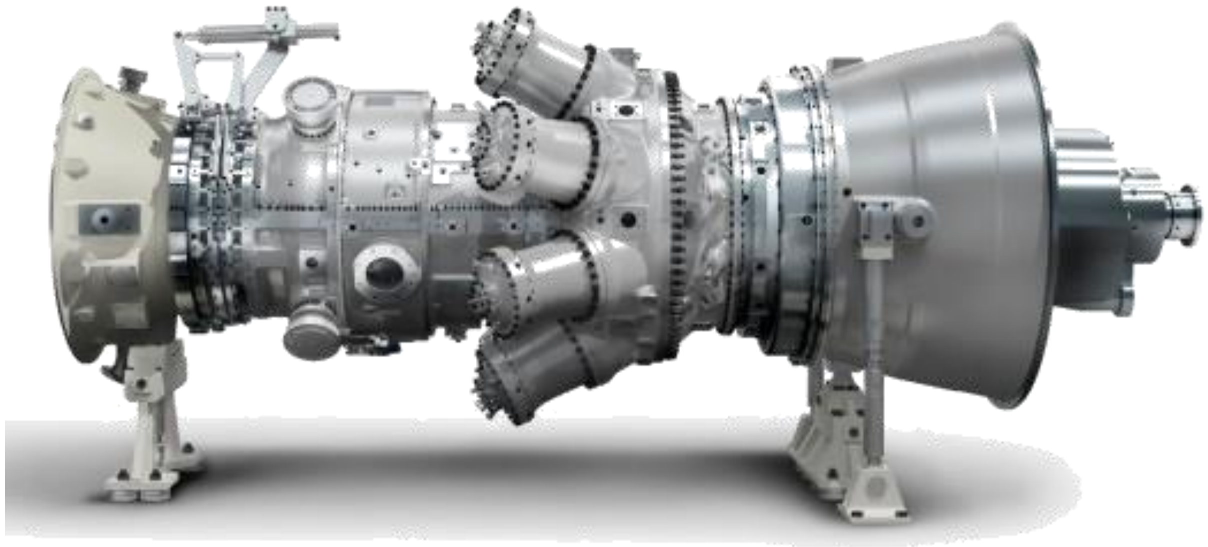
Hardware Issue Development

Current Advanced Pattern Recognition

Great Indicator of Differences

Lacks Causation Indication

90%+ of Alarms Results in Model Retraining



Machine Learning & 1st Principle Physics

Increased Sensitivities to Real/Actionable Issues

Allows for Computer Aided Diagnostics/Prognostics

Requires Data and SME Knowledge

Advanced Pattern Recognition AI

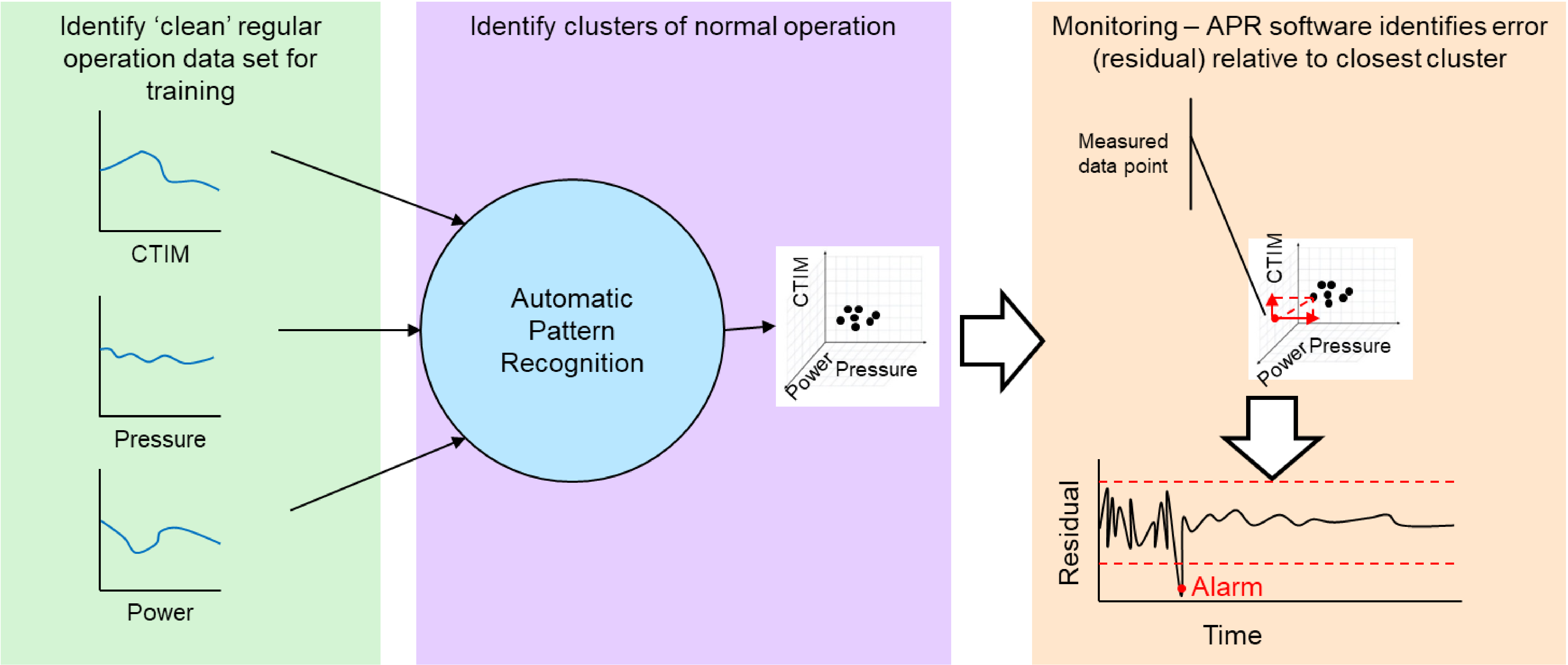
An Example

Let's assume you train your APR on pictures of mallard ducks

APR will know this is a mallard duck, so all is normal!

APR will know it's not a duck, but what is it? Should I be worried?

APR Training – Today (Generalized)



*Note this is a general process description – specifics depend on details of model and software

Implementing AI Requires Solid Fundamentals In Place

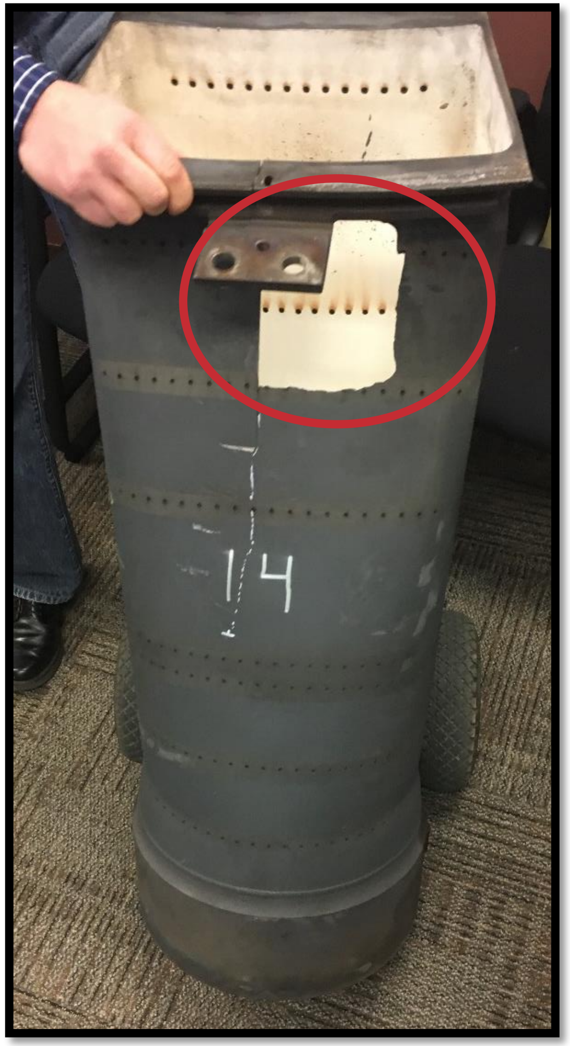
Pillars of Next Generation Monitoring and Diagnostics

Artificial Intelligence seeks to bridge the gaps between these pillars

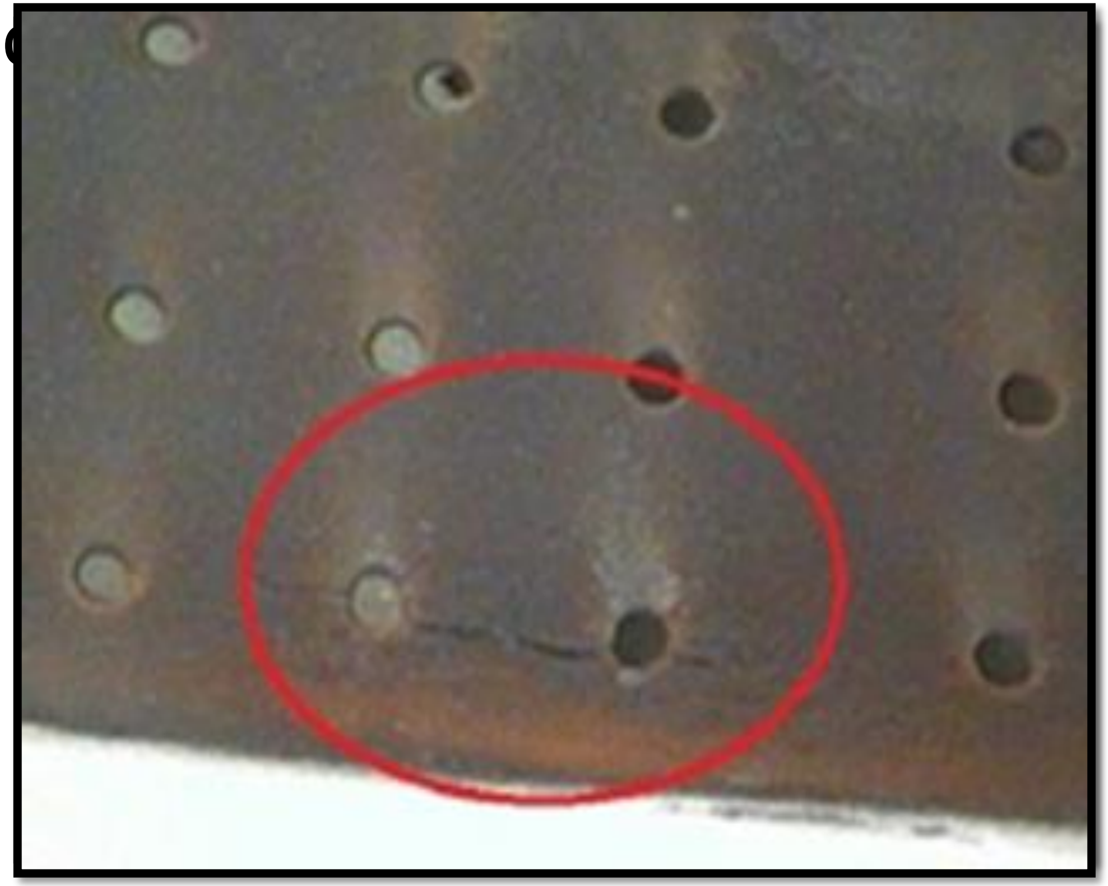
Data	Experience	Modeling
Collection	Grey Beards	Trending
Storage	Empirical Knowledge	Data Driven
Aggregation	Feedback	Physics Modeling
Analysis	Record Keeping	
Use	Domain Experts	

Example of Applying AI to Combustion Hardware

Early Detection of Transition Piece/Liner Failure

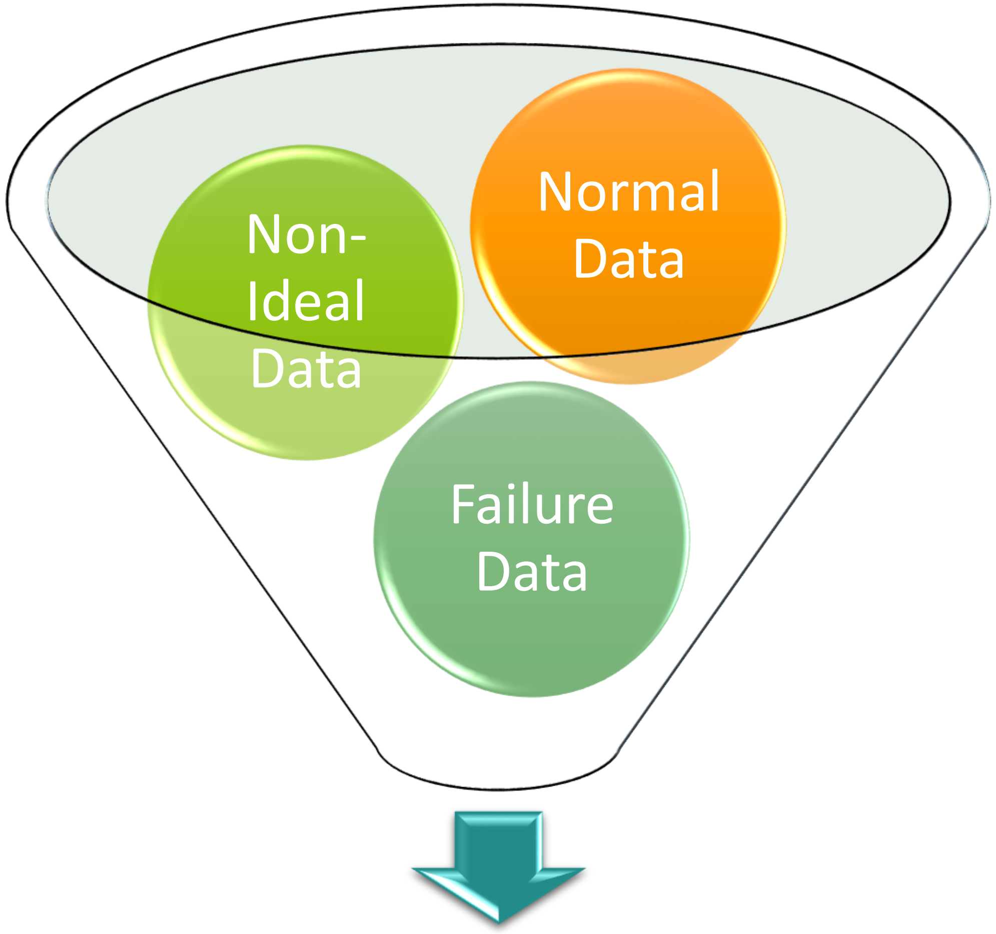


- Detect damage early before forced



Data + Analytics: Information in Context

- Pathway was *enabled* by data
  - Needed normal data
    - 10 different units
    - 6 to 24 months of data each
  - Needed non-ideal data
    - 20 different abnormal data examples
  - Needed failure data
    - Multiple failure types
    - Pivotal for testing analytic capability



Successful Analytics

Fleet Size Implications (U.S.)

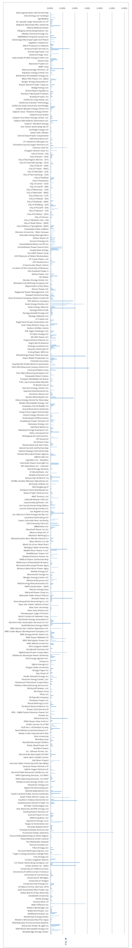
Aviation

- Fleet Size: 6,871
  - But ~2 engines per plane! (~14,000 total)
- Number of operations: 42,270 per day
- Number of operators and respective fleet size:

% of Total Aircraft

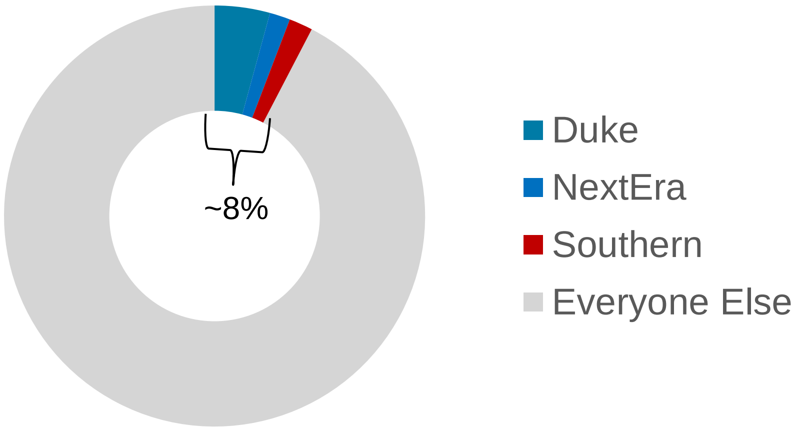
4 operators own 50% of the assets

Power Gen



- 4,100 NG turbines across more than 800 utilities! (as of 2017)
- Some part of larger groups – still data sharing issues due to regulations
- 3 largest U.S. utilities by market cap only own ~8% of NG gas turbines

% Of Total Gas Turbine Fleet



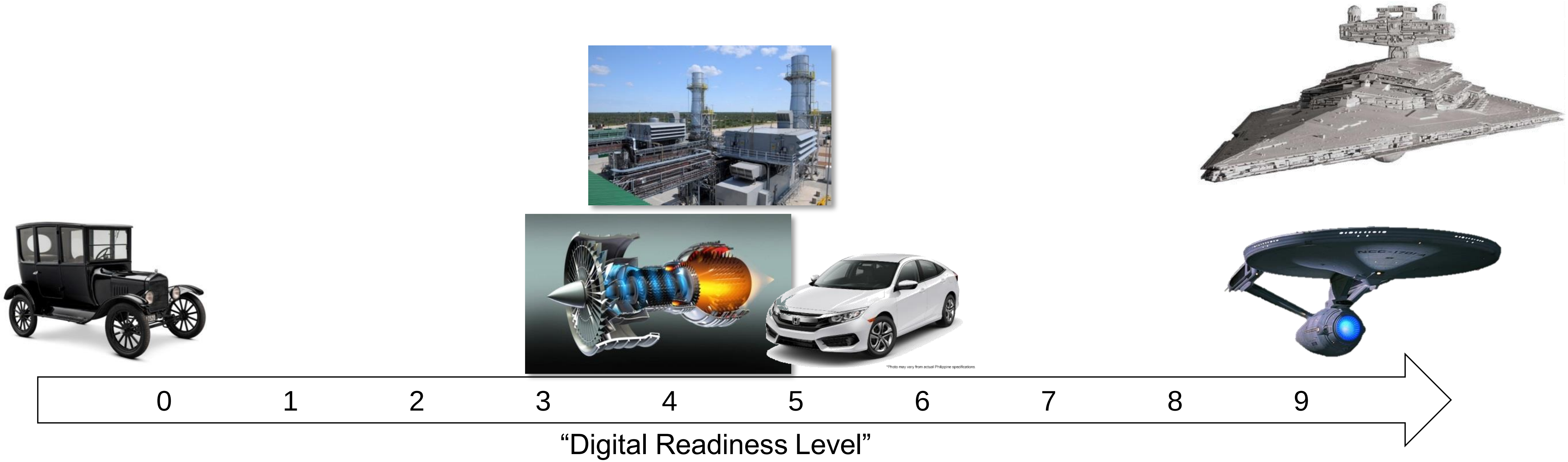
DATA SIGNIFICANTLY MORE CONCENTRATED IN AEROSPACE

What Can Be Learned From Other Industries?

So Where Is Everyone Else?

- No one is close to complete utopian integration of knowledge, modeling, and data for diagnostics

- Where is power-gen compared to aviation sector?
  - Different sets of challenges



Aviation – Advanced Diagnostics Challenges

Data

787 generates 500 GB of data per flight

Much more than just the engine to monitor

- Airframe
- Avionics
- Sensing Equipment
- Etc…

$100 / MB using legacy in-flight data links

Sensor availability - GE helped with digital twin of landing gear – by adding additional sensors

Experience

Going back to basics – combining 50 disparate IT systems through central data repository team

One Million+ handwritten defect reports

Paper accounts for 90% of airline records

Models (and Analysis of Data)

Airline Success Stories

GE Digital Twin – Landing Gear

Additional sensors added to common failure points to correlate to operations

GE90 Fan Blade

After observing distress – correlated operating profile to suggested maintenance action

Lufthansa Altitude Sensor Errors

Analyzed data of autopilot failures to develop bad altitude sensor early detection

Delta Operational Efficiency

Maintenance related cancellation drops from 5,212 to 123 in 6 years

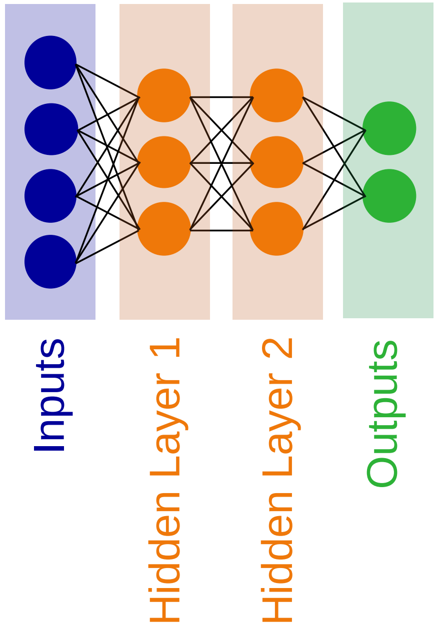
- Noteworthy successes, but all are post issue, diagnostics
- Large data sets aid in analysis
- There may be predictive catches, but aren’t being publically touted yet

- Establishes tolerance ‘bands’ for every part
- Part is pulled EVERY TIME measurements exceed band
- Rigorous inspection used to update tolerance – adjust limits
- Uses Smart Signal…

Quick Summary of Archetypes of Machine Learning

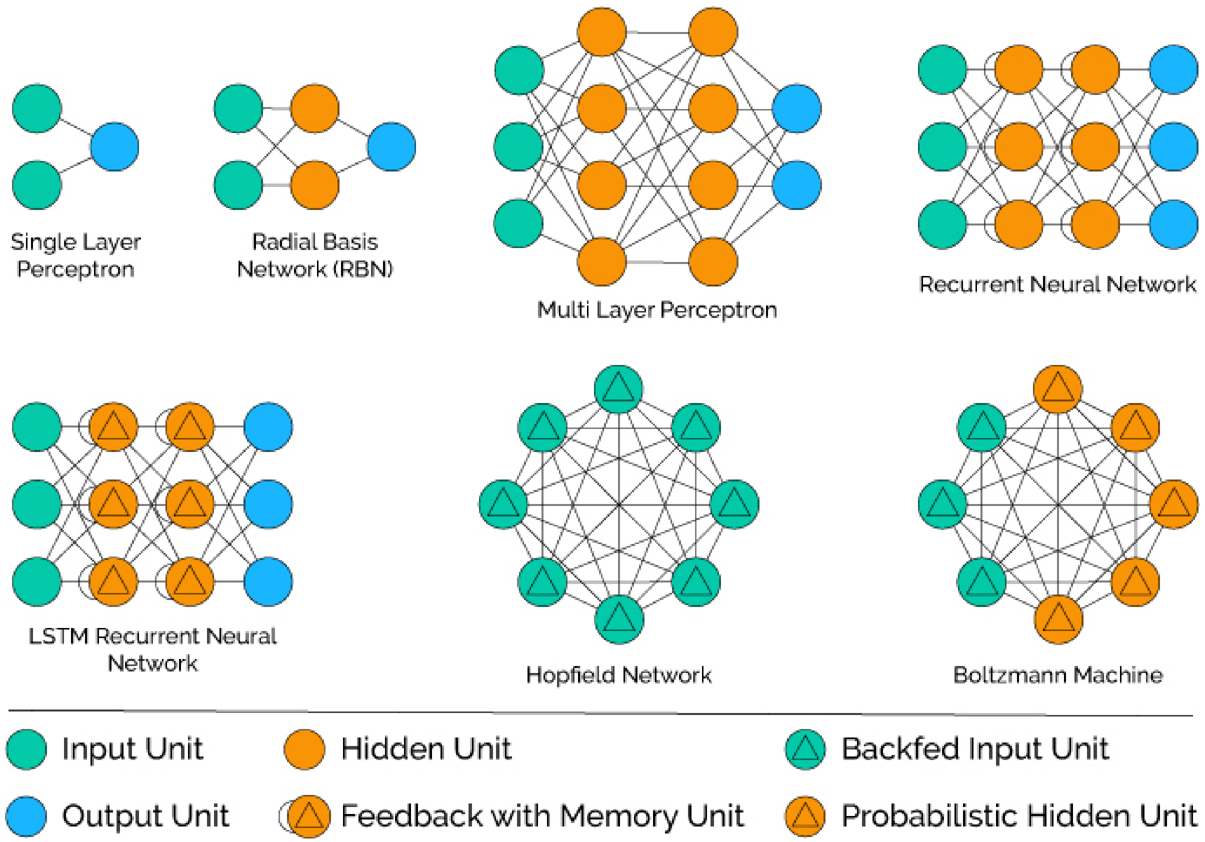
Artificial Neural Networks – What are they?

- Designed to mimic the connection of neurons in the human brain
- Nominally consists of 3-4 layers (multi layer perceptron)
  - Input layer
  - One to two neuron layers (hidden nodes)
  - Output layer
- Both deterministic and probabilistic types exist
- Static and 'learning' or updating models exist



Artificial Neural Network - Uses

- Uses
  - Fitting models to observed data
  - Fitting models to computer generated data
  - Classification
- Types

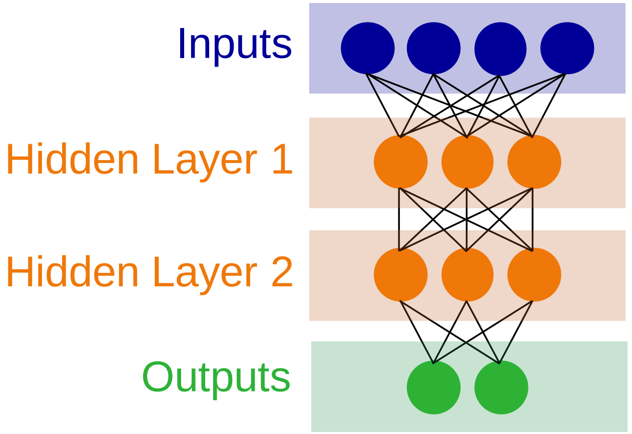


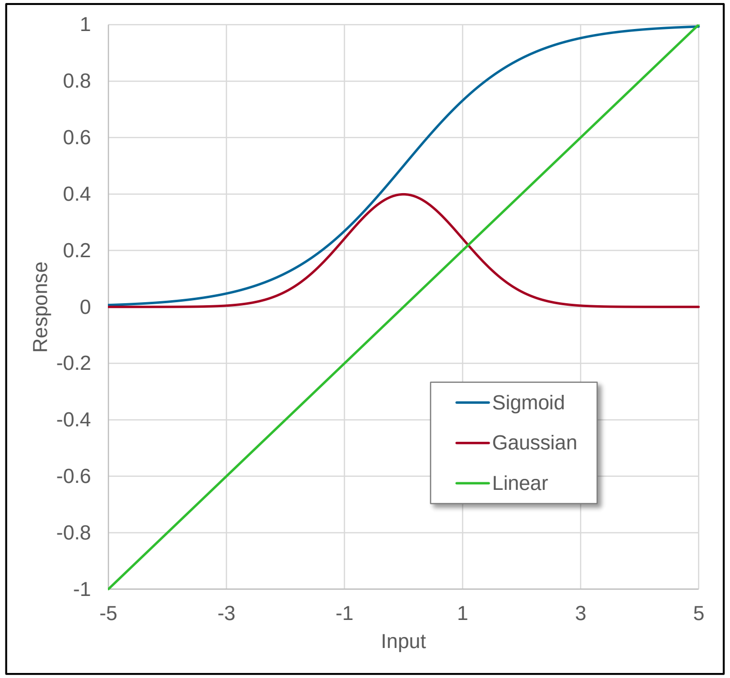
Artificial Neural Network - Uses

- **Pros**
  - Can adapt to discrete and non-linear responses
  - Computationally efficient and portable once trained
  - Can handle both discrete and continuous inputs simultaneously
- **Cons**
  - Easy to over-fit
  - Can require more extensive data set for training
  - Can be guess and check on network structure

Artificial Neural Networks – Common Functional Forms

- Input layer: Regression variables
- Hidden Layers contain activation functions
- Hidden Layers
  (commonly one or two)
  - Sigmoid $f(x) = \frac{1}{1+e^{-x}}$
  - Gaussian $f(x) = e^{-x^2}$
  - Linear $f(x) = x$
  - ArcTan $f(x) = \text{atan}(x)$
  - Other variations, but all have similar characteristics shapes
- Output Layer
  - Linear combination of last hidden layer
  - $Y = aH1(bx + c) + eH2(fx + g) + \cdots$
- Backpropagation algorithm solves for coefficients





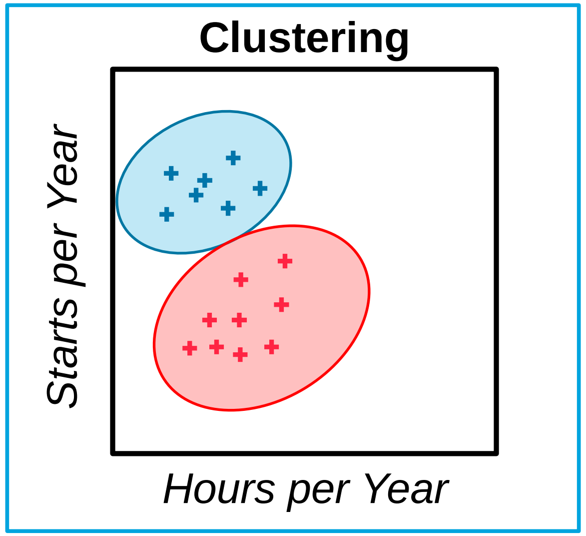
Typical Activation Functions

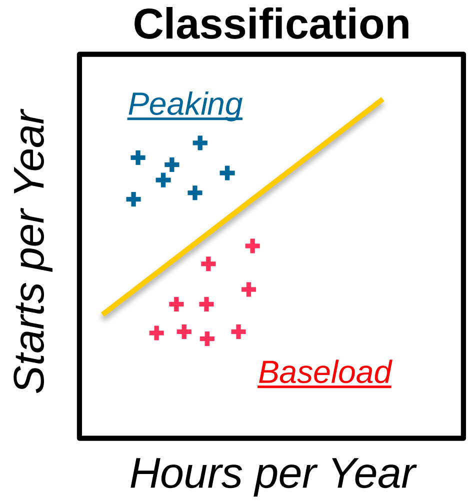
ANN Major Considerations - Overfitting

- Neural networks are more complex
- Overfitting can lead to erratic behavior
- Provides inconsistent predictions away from training points
- Can cause issues if used in numerical simulation (including APR)
  - Most models work better if underlying functions are smooth with slowly changing gradient
  - Fortunately most engineering problems are also 1st or 2nd order
- Another reason training data quality is critical
  - A neural network can fit the data if given enough degrees of freedom

Clustering Algorithms - Uses

- Uses
  - Unsupervised learning
- Common Types
  - K-Means
  - Hierarchical
  - Normal Mixtures
- **Pros**
  - Useful when functional form of data is not known or hard to define (does not mean it does not exist!)
  - Easy to use and understand
- **Cons**
  - Lack good ability to extrapolate
  - Choosing the number of clusters can be difficult
  - Geometrically based!
  - Dependent on magnitude of data if data not normalized

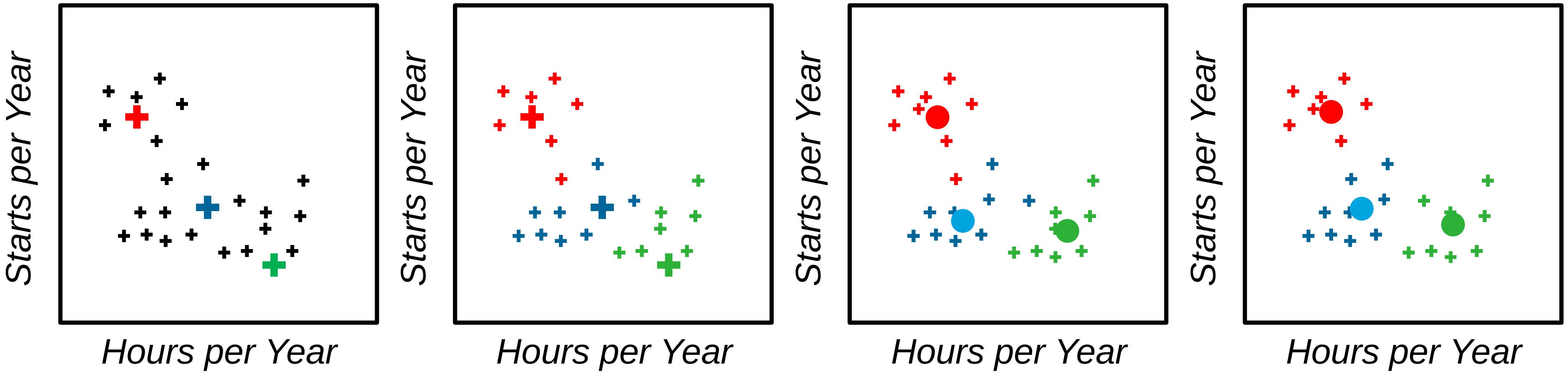




Clustering Types: K- means

- One of the more common types is called k-means clustering
- Forms clusters on k (user selected) means in the dataset
- As an example define boundaries for peaking, cycling, and baseload operation based solely on data

<u>K-Means Clustering Process (3 clusters)</u>



Step 1: Randomly pick three 'means'

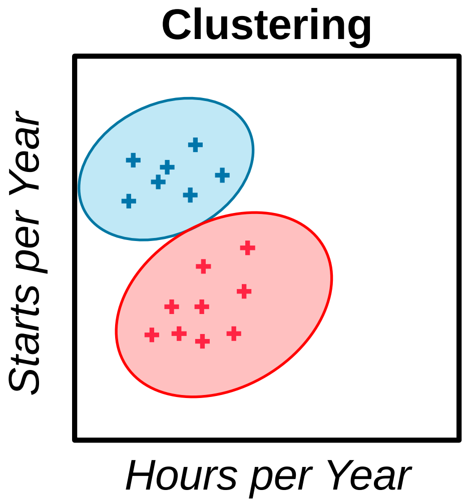
Step 2: Assign membership based on Euclidian distance

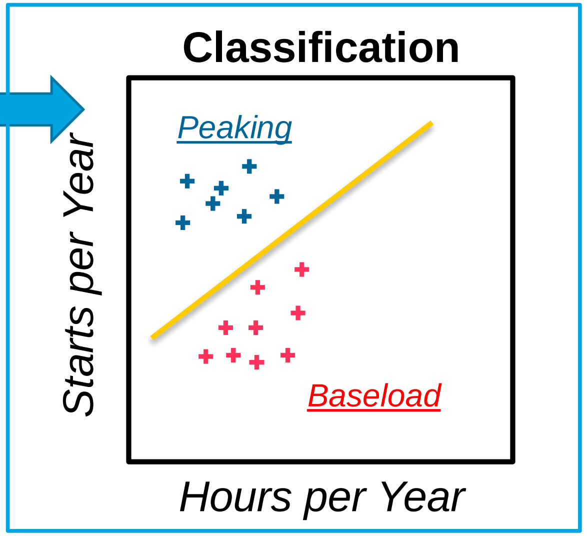
Step 3: Calculate centroid (c.g.) of each cluster

Step 4: Re-cluster on centroid – repeat 2 – 4 until converged

Classification Algorithms – Uses

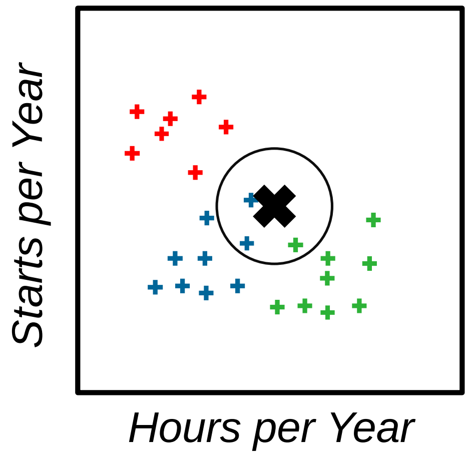
- Uses
  - Appropriate when training dataset already 'tagged'
- Types
  - Logistic Regression
  - Naïve Bayes Classifier
  - K-Nearest Neighbors
  - Decision Trees
  - Neural Networks
- **Pros**
  - Several options available
  - Conceptually easy to understand
  - More complex functional forms available
- **Cons**
  - Relies upon prior knowledge of group membership
  - Some are geometrically based





Classification Example Applications – K Nearest Neighbors

- Similar to clustering approach, except response is the average of the k-nearest neighbors
- For a new point – finds k nearest neighbors
- Largest number of matches yields class association
- Choosing the right k is trial and error



- Assume k set to three
- New point at **X**
- 3 nearest neighbors are two blue and one green
- Membership is blue

Bayesian Networks – What Are They?

- Follows Bayes Theorem:
  - $P(Y|F) = \frac{P(F|Y)P(Y)}{P(F)}$
- The power behind Bayesian Networks lie in the fact that:
  - Prior beliefs can influence posterior (future) thinking based on new observations
  - Allow for model to learn over time as new data becomes available
  - Probabilistic

Bayesian Networks – A Simple Example

Have Kids?	Yes:	No:

Over 40?	Yes:	No:

P(kids) = ?

P(over 40) = ?

$$P(Over\ 40|Have\ Kids) = \frac{P(Have\ Kids|Over\ 40)P(Over\ 40)}{P(Have\ Kids)}$$

Bayesian Networks – A More “Real World” Example

P(immun syst)
0.05

P(smoking)
0.3

P(common cold)
0.35

P(lung cancer	smoking)
0.1	true
0.01	false

P(bronchitis	smoking)
0.3	true
0.01	false

P(runny nose	common cold)
0.9	true
0.01	false

P(pneumonia	immun syst,	lung cancer)
0.3	true	true
0.3	true	false
0.05	false	true
0.001	false	false

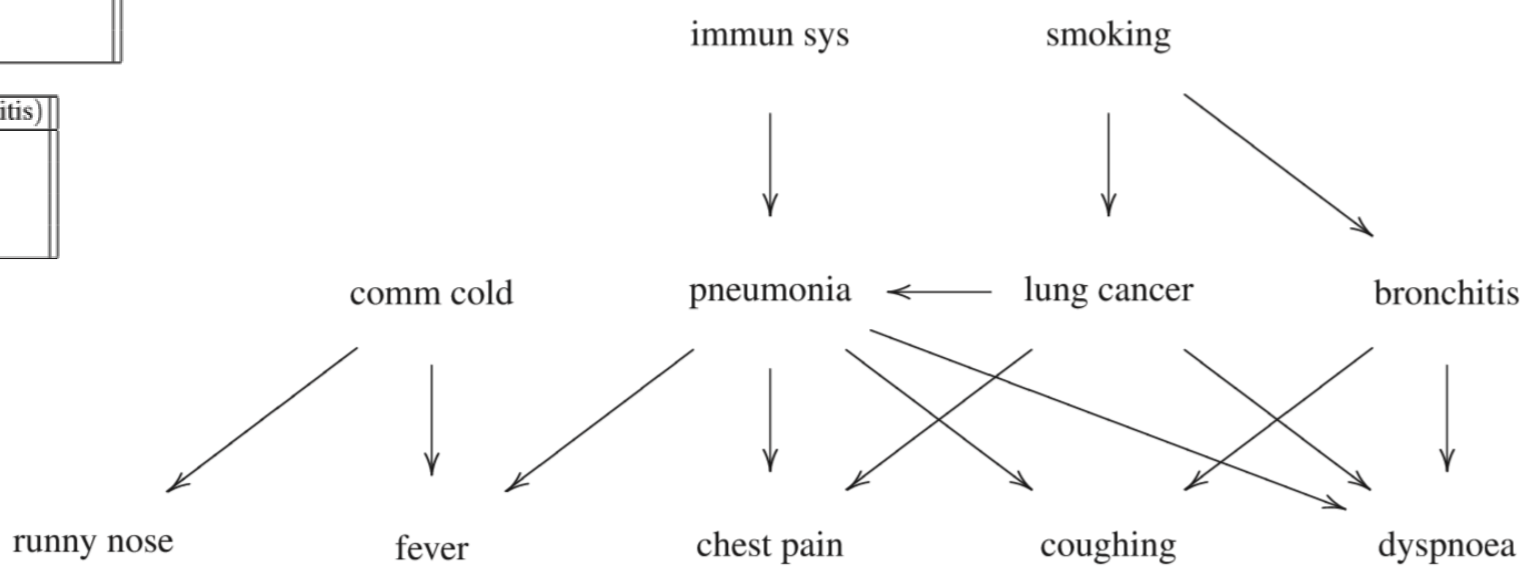
P(fever	pneumonia,	common cold)
0.9	true	true
0.9	true	false
0.2	false	true
0.01	false	false

P(cough	pneumonia,	bronchitis)
0.9	true	true
0.9	true	false
0.9	false	true
0.1	false	false

P(chest pain	pneumonia,	bronchitis)
0.9	true	true
0.9	true	false
0.9	false	true
0.1	false	false

P(dyspnoea	bronchitis,	lung cancer,	pneumonia)
0.8	true	true	true
0.8	true	true	false
0.8	true	false	true
0.8	true	false	false
0.5	false	true	true
0.5	false	true	false
0.5	false	false	true
0.1	false	false	false

Add layers to map to GT



*Wiegerinick, W., Burgers, W., Kappen, B., “Bayesian Networks, Introduction and Practical Applications”

Bayesian Learning – Uses

- Uses
  - Model calibration
  - Diagnostics
  - Model Updating
- Pros
  - Flexible
  - Can learn over time
  - Suitable for discrete and continuous data
  - Good for mixed data sets
- Cons
  - Often difficult to setup
  - Validation tricky
  - Often requires coupling with additional modeling (i.e., neural networks)

Idea Behind Bayesian Calibration

- Use assumed prior belief coupled with observations to update your prior belief
- Also takes into account measurement and model representation error
  - Model representation error known from regression (prior slide)
  - Measurement error can be assumed based on sensor types
- All values are really distributions
  - Conceptually think of every measurement & prediction as having a +/- intrinsically associated with it

'Unknown' Measurements (i.e., Burner dP)

Model Representation Error

$$y = f(x_1, x_2, \dots, h_1, h_2, \dots) + \delta + \varepsilon$$

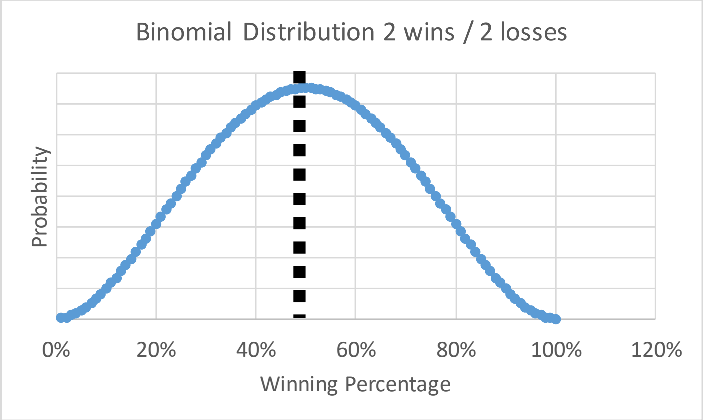
'Output' of model (i.e., power)

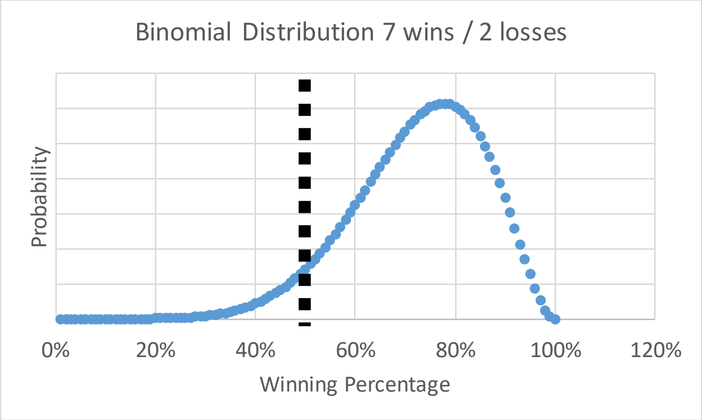
'Input' Measurements (i.e., CTIM)

Measurement Error

A Simple Example – Winning Percentage

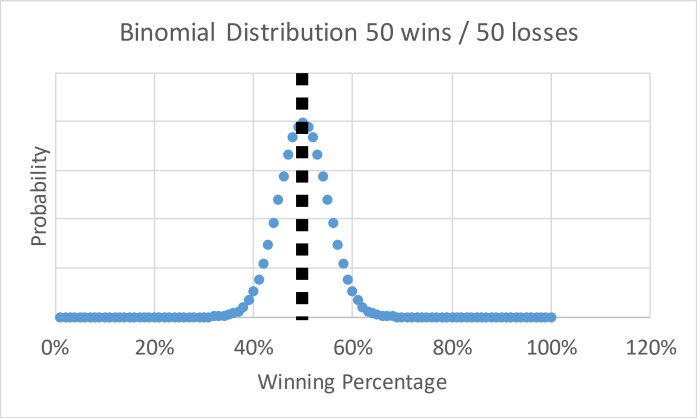
- A binomial distribution shows expected win rate
  - Useful for example since it is a ‘closed form’ update
- Example 1: Little prior knowledge
  - Let’s assume I know my favorite team has 2 wins and 2 losses
  - The winning percentage is 50%, but how sure am I that is the true value?
  - This curve represents my prior belief
  - Looking at the spread it says I’m open to changing my opinion
- Let’s say my team goes on to win 5 in a row (so they are now 7 and 2)
  - Now I’m fairly convinced they are an above 50% team
  - Still some uncertainty as to how much better

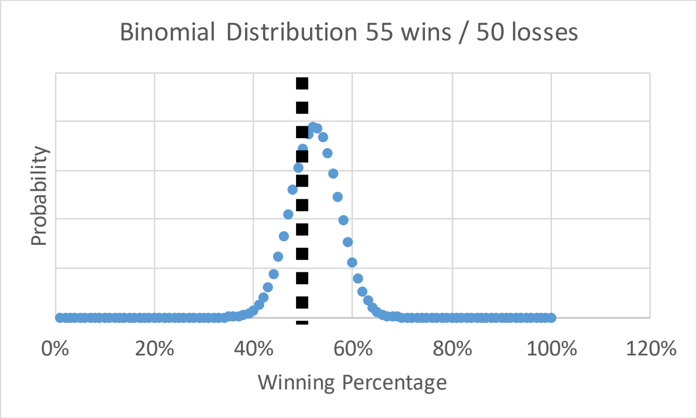




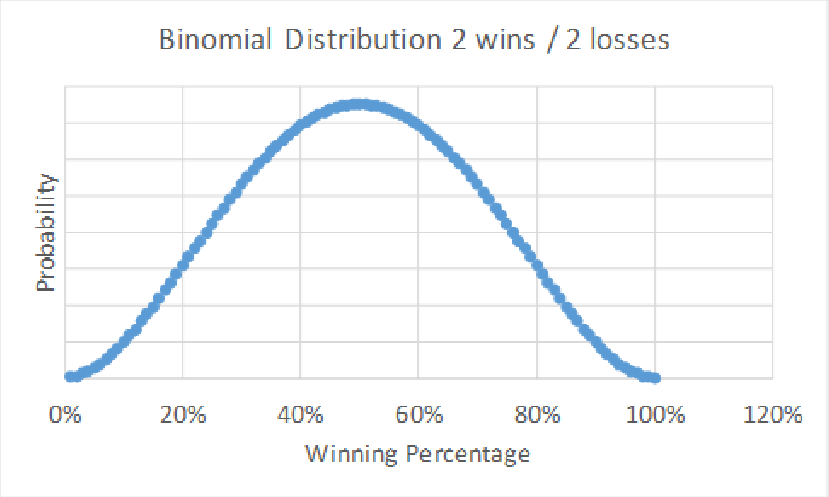
A Simple Example – Winning Percentage - Continued

- Now let’s assume my prior knowledge is that the team has 50 wins and 50 losses
  - Same winning percentage (50%) as prior example
  - More evidence, so I’m more certain
- Assume the team wins the next 5 games, same as before
  - Now 55 wins and 50 losses
  - Still shits my opinion, but the meat of my opinion is that they’re still close to a .500 team

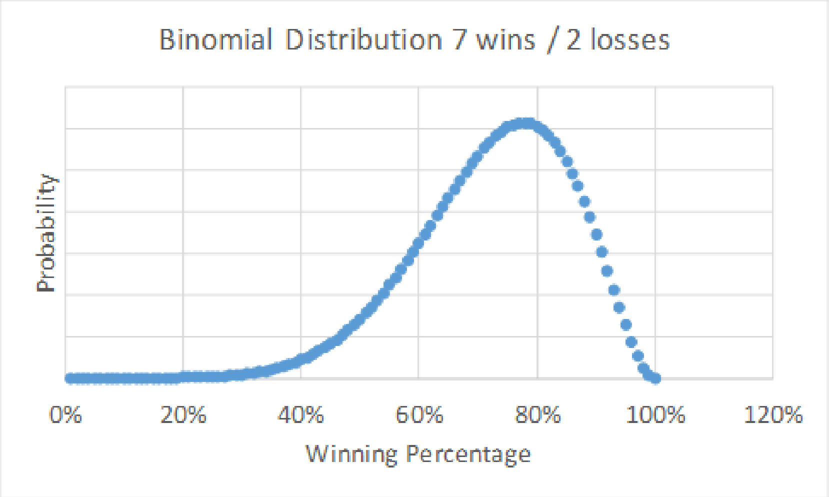




Winning Percentage – Putting into Bayesian Speak



5 game winning streak



Prior Belief + **New Observations** = **Posterior Knowledge**

Gas turbine model more complex, but same basic idea:

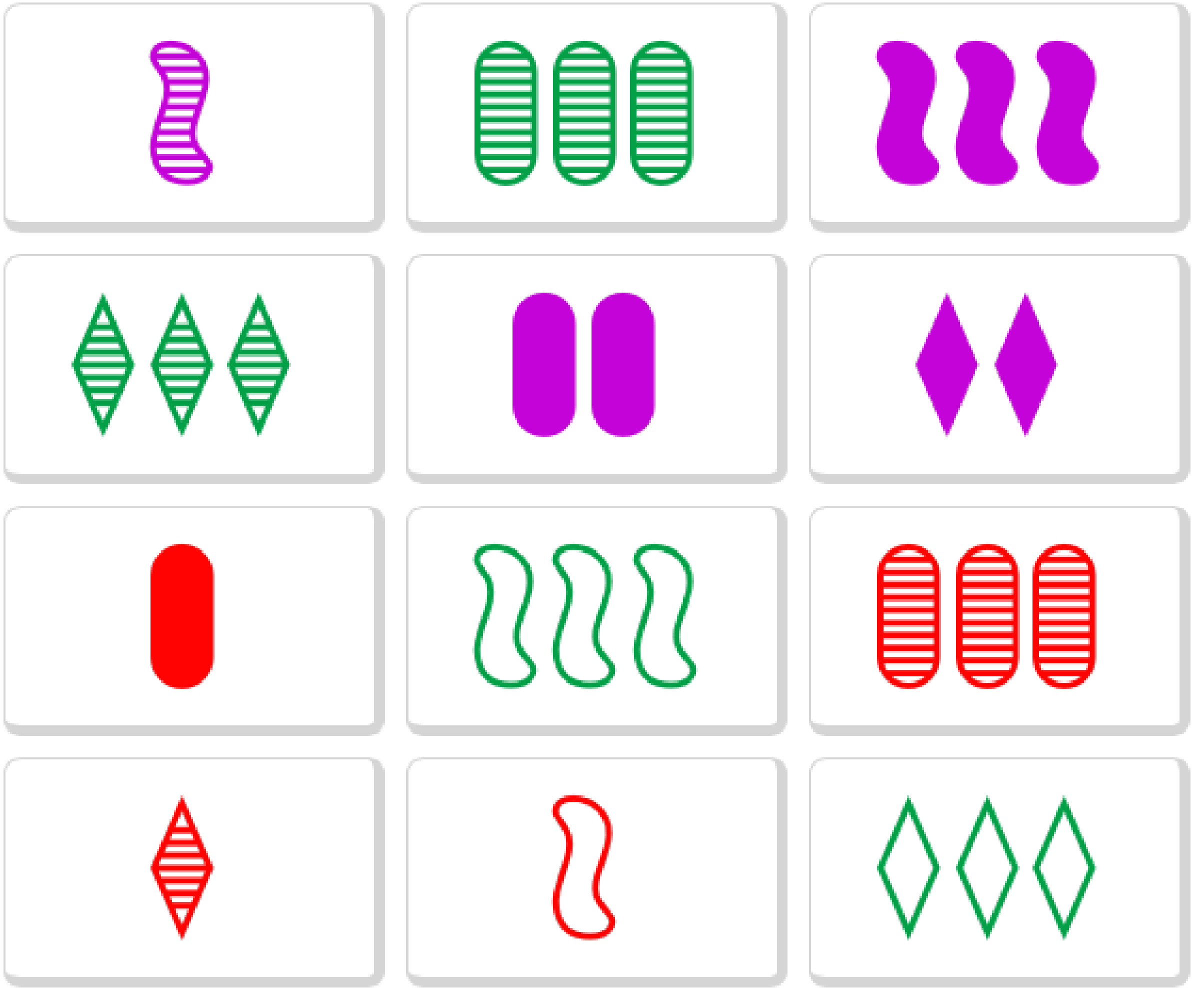
There are health and performance parameters which influence the performance of the machine – we want to estimate them based on our working knowledge of the hardware

Examples and Relevant Impacts

Where are (many of us) today?

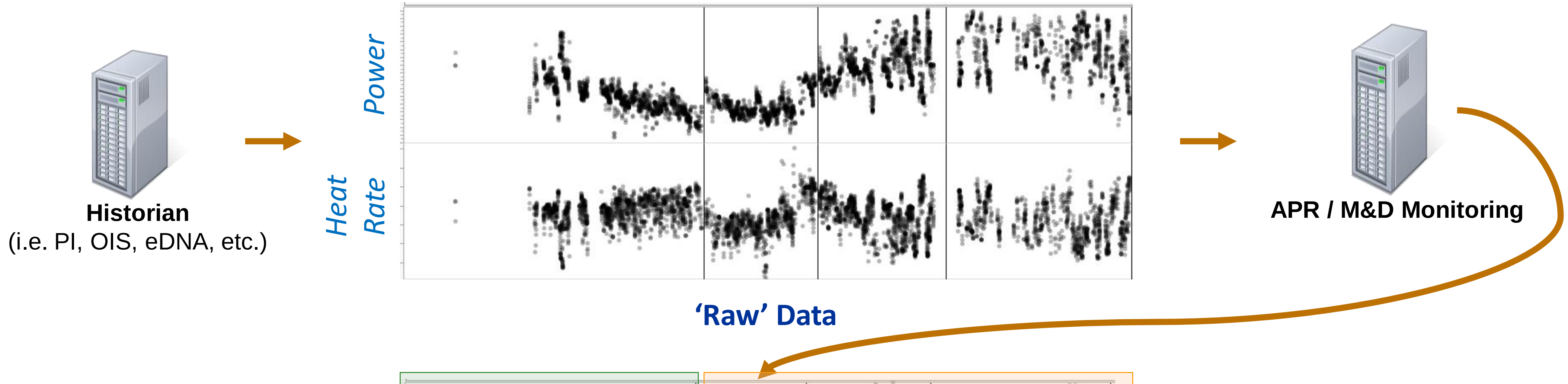
- Many are using advanced pattern recognition
- Strengths:
  - Generalized (Unsupervised)
  - Easy to use – minimal engineer training required
- Weaknesses:
  - Does not provide right level of **interpretation** for *complex* problems

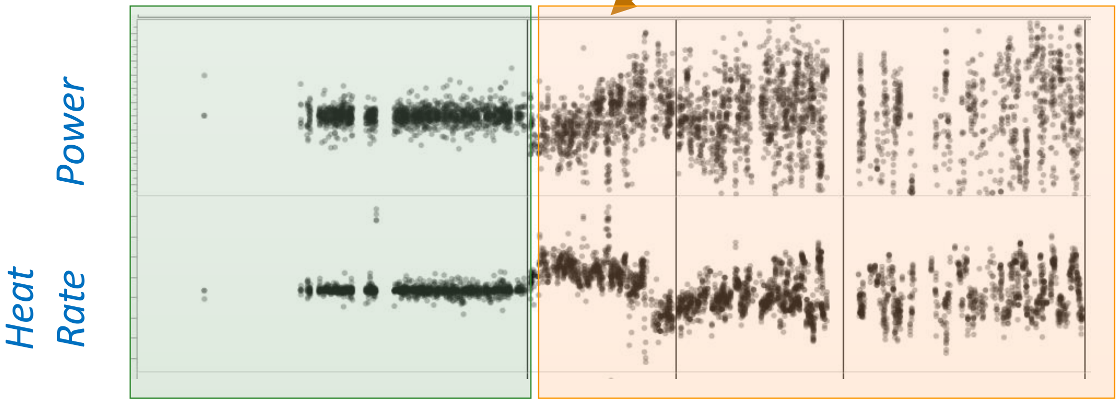
How would you identify patterns at right??



Where are we going???

- Generalized APR is focused on gross changes and is often correlation based



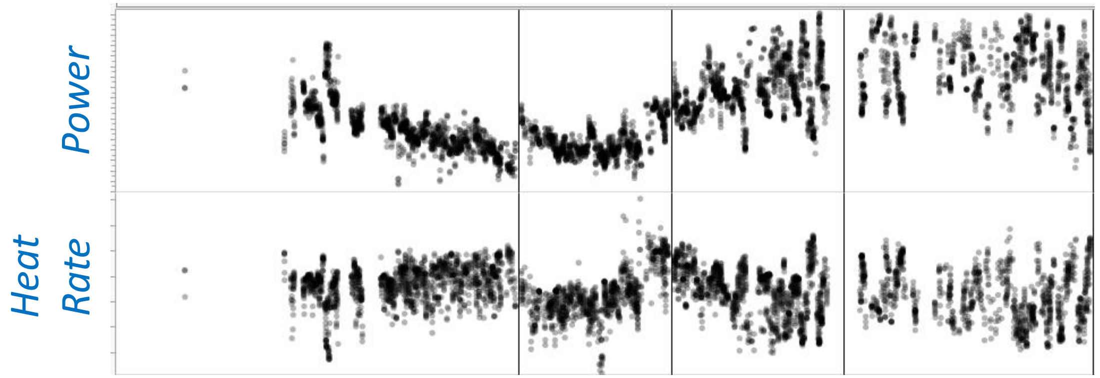


- APR good for gross error in training set
- Error grows as time goes on – spend all day tracking false alarms

Where are we going???

- Don’t let the APR do the correlating!
- Insert a physical or expert model into the process

Historian
(i.e. PI, OIS, eDNA, etc.)



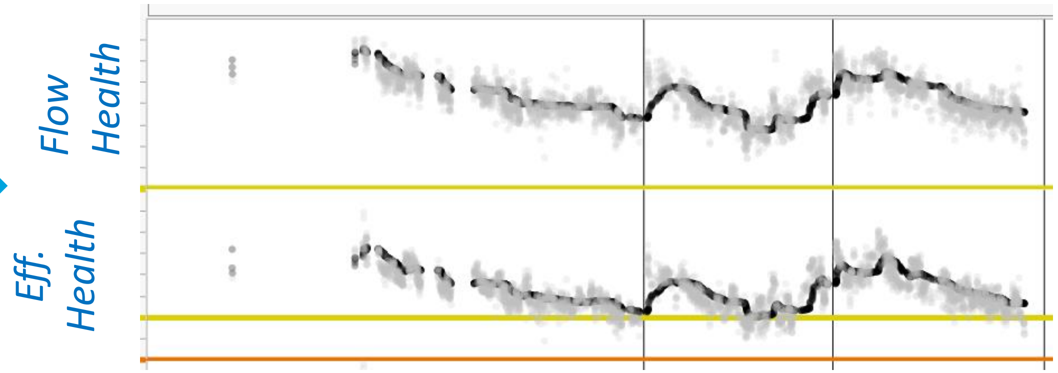
‘Raw’ Data

Digital Twin

Physics Model

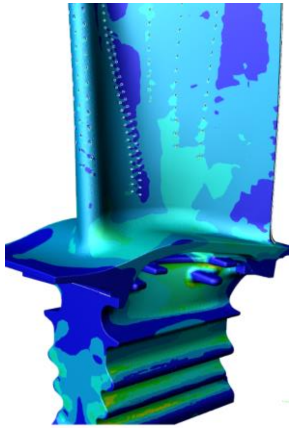
Expert System

Virtual Sensors

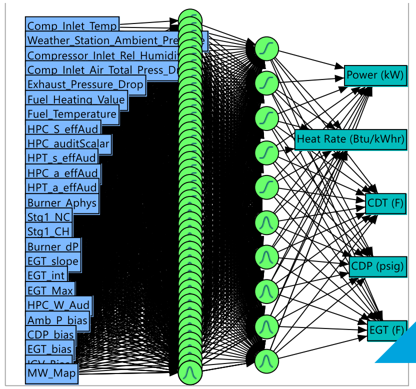


- Use physical model to provide ***insight*** into raw data
- Can still use APR or other techniques for trending and error detection

Ways to incorporate higher fidelity



Physical Model



Encapsulation through AI/ML

Digital Twin

Physics Model

Expert System

Digital twin

Issues and Challenges That Must Be Addressed

Large Challenges Remaining

- **Data and Data Management**
  - "Intentional Data"- data is available for performance and reliability
  - Standardized clearing house for diagnostic and inspection data
  - Any part / any engine / any time / immediately available

- **Experience**
  - Standardized inspection reports and information
  - Digital reports
  - Automated correlation with monitoring tools

- **Modeling**
  - Seamlessly merging experience, AI, and physics
  - Usability of high fidelity tools by the masses
    - Consider the iPhone vs. first IBM mainframe
    - AI helps enable this!
  - Intimate and automated feedback between modeling and data collection

Together…Shaping the Future of Electricity