Practical Machine Learning Applications

CIGRE Grid of the Future 2021

T. Rhodes, Duke Energy, USA
I. Mitiche, Glasgow Caledonian, UK
F. Werner, Doble Engineering, USA
P. Boreham, Doble Engineering, UK
Tony McGrail, Doble Engineering, USA

Overview/Introduction

- Overview/Introduction
- Background
- Health and Risk Management at Duke Energy
- Some issues with Giraffes
- Out of Distribution and ML Classifications
- Discussion

Presenter Background

- National Grid UK: substation tech specialist, transformers
  - Go/NoGo decisions: timescales & actions
- National Grid US: Substation Asset Mgr.
  - >2,000 power transformers, many >80 years old…
- Doble Engineering
  - Asset management and Monitoring Technology

Industry 4.0: Technical factors…

INFLUENCING FACTORS
(Market Trends and Dynamics)

MARKET DRIVERS:

- Development of a secure cloud computing platform
- Increased use of wireless communication technology for condition monitoring of equipment
- Inclination of end users toward predictive maintenance
- Adoption of automated condition monitoring technologies

MARKET RESTRAINTS:

- Lack of trained technical resources for data analysis
- Lack of trust in prediction capabilities of machine condition monitoring technology

MARKET OPPORTUNITIES:

- Advent of Big Data analytics
- Strategic partnership and collaboration

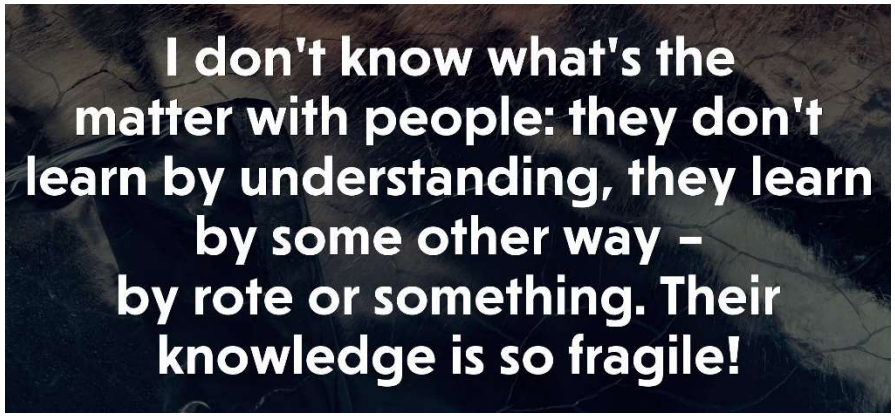
MARKET CHALLENGES:

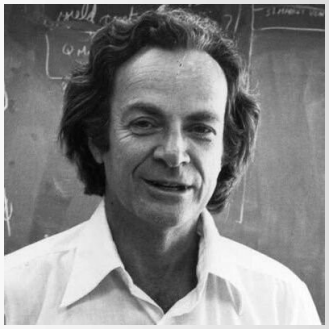
- Incurrence of additional expenses and complexities in retrofitting existing systems
- Availability of expertise at remote locations

- **ACQUISITIONS, EXPANSIONS, & PARTNERSHIPS:** More than 10 deals in the last 3 years
- **PRODUCT LAUNCHES:** More than 25 product launches in the last 3 years

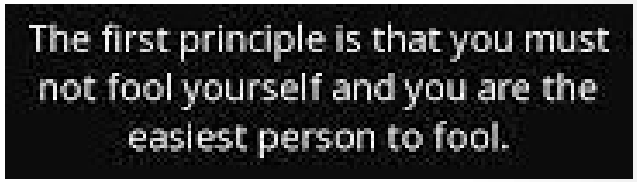
Richard Feynman & Others: Quotations

"…you have to have an understanding of the connection of the words with the real world."





I hate quotations. Tell me what you know.
Ralph Waldo Emerson



"There is a notion of success ... which I think is novel in the history of science. It interprets success as approximating unanalyzed data."
Noam Chomsky

Duke Energy: HRM Program

WHY have a Health and Risk Management AI solution?

WHY we need Artificial Intelligence to make it work?

- 10,000 Large Power Transformers (Banks > 7.5 MVA)
- 25,000 Substation Circuit Breakers
- Dozens of Data Sources
- Millions of Data Points
- Decreasing O&M dollars
- Decreasing number of people to evaluate and manage all the data, and retiring most experienced personnel

HRM Targeted Benefits

- Manage replacement of aging fleet
  - Reduction in Emergent/Emergency Replacements
- Expectations are to do less maintenance but extend asset life
  - Avoided Cost (Good Catches)
  - Condition Based PM Schedules
  - Extended Diagnostic PM Schedules w/ CBM
- Reduce Customer Minutes Interrupted from Transmission Asset Failures
- *Operational Efficiencies*
- Capture Expertise of SME's

Earlier Attempts at ML for HRM Failed at Duke Energy:

"We tried it, Machine Learning doesn't work on Transformers"

~ Un-named Duke Engineer 2017

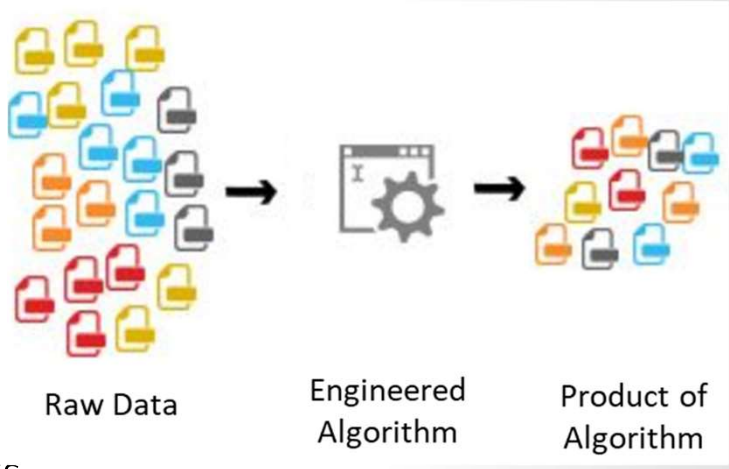
- Limited and Bad Data
  - Failure to document and maintain failed asset data
  - No investment in cleaning and verifying data we did have
  - Data not normalized across multiple sources nor within a single source
  - Unique characteristics of hand made large transformers
- Depended on IT and ML professionals to develop methods and analysis not Transformer SME's
- Most ML is Gaussian
  - Failure mode for Transformers is non-Gaussian
    - Doesn't follow linear progression
    - Doesn't follow the Bathtub curve

"Your data must be wrong because if we use it our models don't work"

~ Un-named MIT Grad & XX Data Scientist 2019

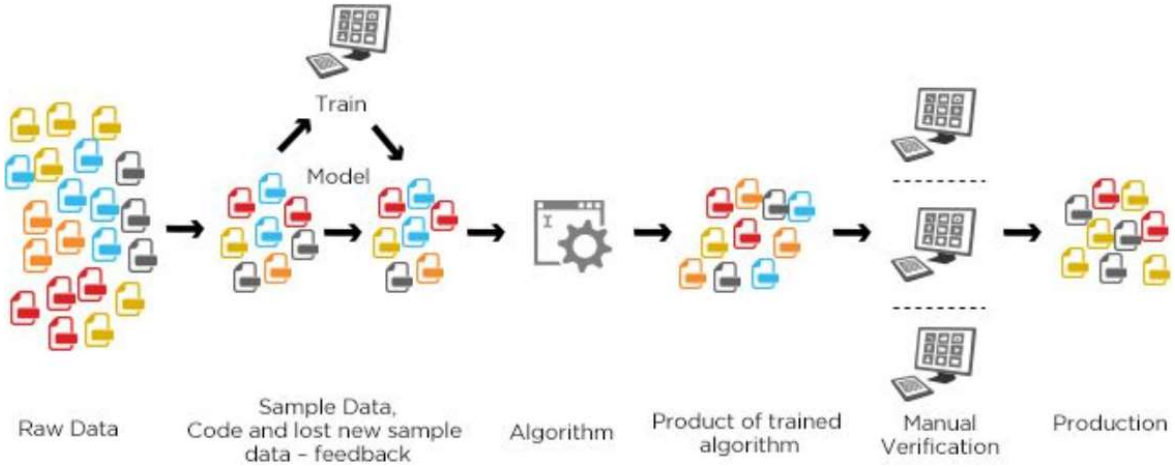
Engineering Analytics

- Only as good as the calculations are designed
  - With nearly a century of analysis, the engineered analytics have proven track record. (– IEEE C57.104 no longer uses TDCG)
- Only analyzes what it is told to look at
- Generally not Systematic
- Examples:
  - Doble Frank Scores
  - EPRI Normal Degradation Index
  - TOA4 DGA scores
  - Duval Triangle
  - Your Company's Engineers' Spreadsheets

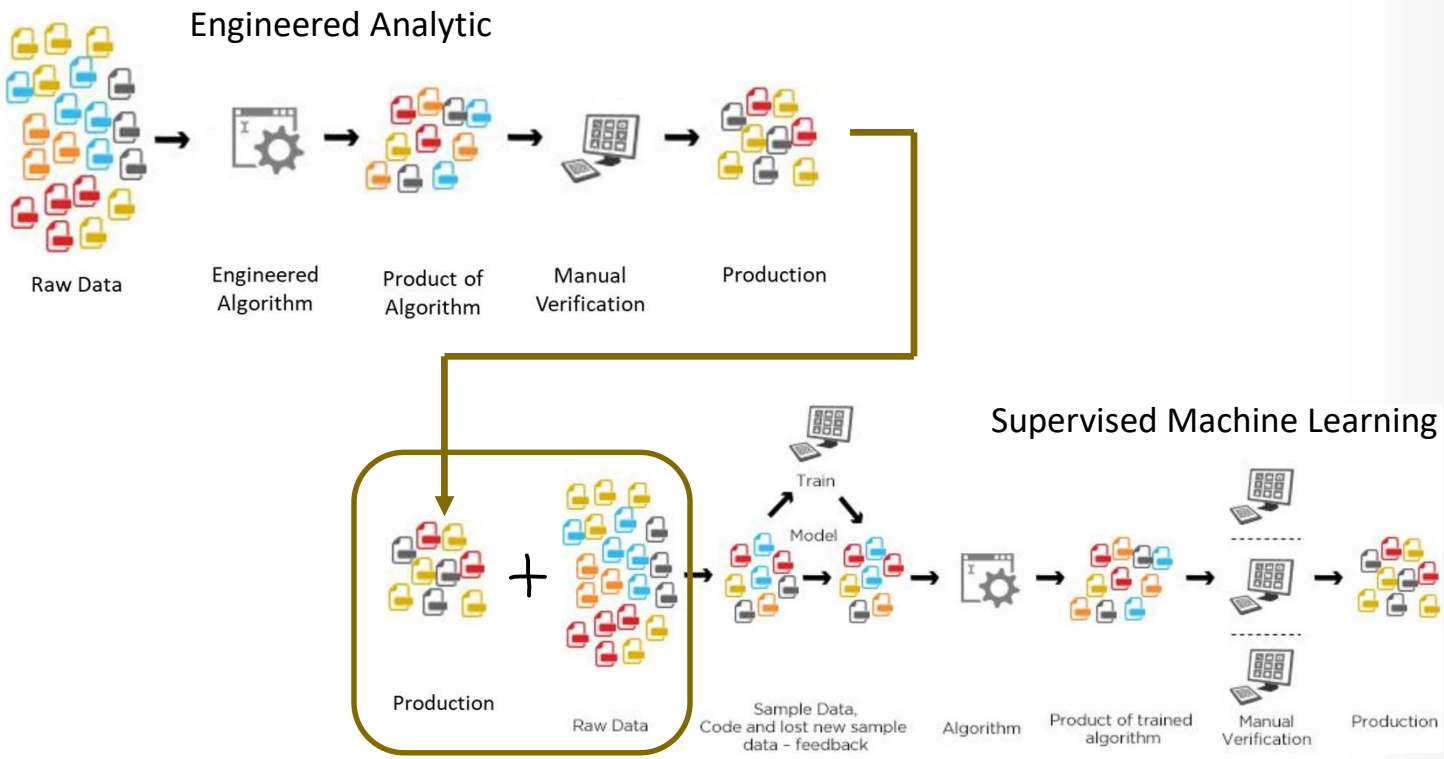


Supervised Machine Learning

- Needs large amount of Data
- Needs known results for subset of data in order to train algorithms
- Uses SME's who understand the data to evaluate features



Hybrid – Engineered Analytics & Supervised ML



Lessons Learned So Far

- Hybrid Model is accurate in predicting and ranking
- Data Cleanup and Maintenance is Essential
  - Initially 80% of our time and effort is spent cleaning and verifying data, and we expect after full roll out 60% of our time will be maintaining data
  - Permanent people to do “data governance”
- You Will have to modify and redo your work.
  - Sustain the continuous improvement cycle
  - Does your budget include time and money to redo work you did (or missed) this year, next year?

Tom Rhodes, Duke/Progress

- Our Secret Sauce at Duke, is not a secret…
  - No Black Box analytics where the SME's can't see exactly how all results are calculated
  - State of an asset to say Replace or Stable must be set by an SME not the system
  - Only automated States are:
    - Monitor (where it sees a WO created to do more testing)
    - Service (where it sees new corrective WO)
    - Risk Identified (trying to get an SME's attention and look at what might be a problem).
- Our biggest problem is the math can show more significance than really exists.
- SME's get it but management this week is asking why a transformer with a 2.678 health score failed before one with a 2.714 health score?
- This ignores the fact that health score is no more accurate to 3 places than a DGA gas reading 130ppm is meaningfully different than 140ppm.
- Managers are generally no longer engineers much less scientists. Our biggest problem is managing what the HRM analytics truly mean, and keep pointing people back to the SME's and teaching that the HRM tool is point SME's at where to look, not make decisions.

"Life is under no obligation to give us what we expect”

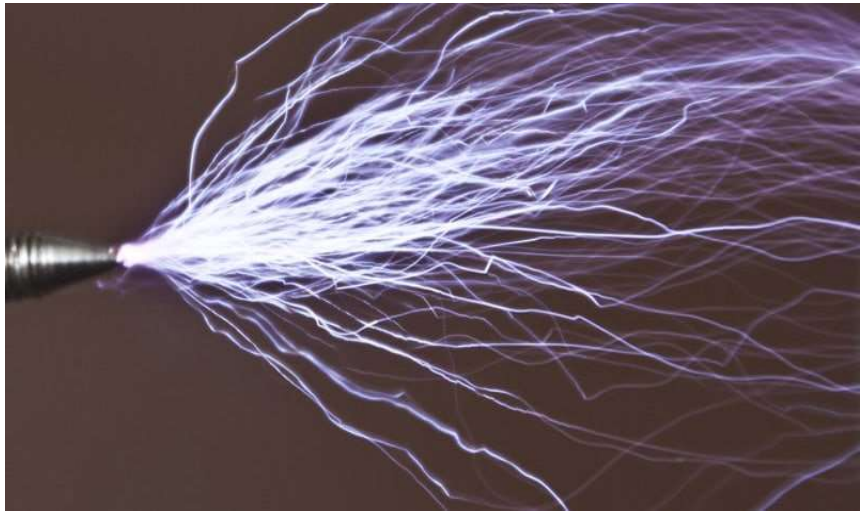
Giraffing

- The occurrence of uncommon exemplars in the training data exceeds their presence in the real world
- Named for the detection of giraffes in photos where, as far as we can tell, no giraffes exist
- The four pictures here were deemed to have giraffes present by Microsoft's state of the art AI tool
- What gave it that idea?

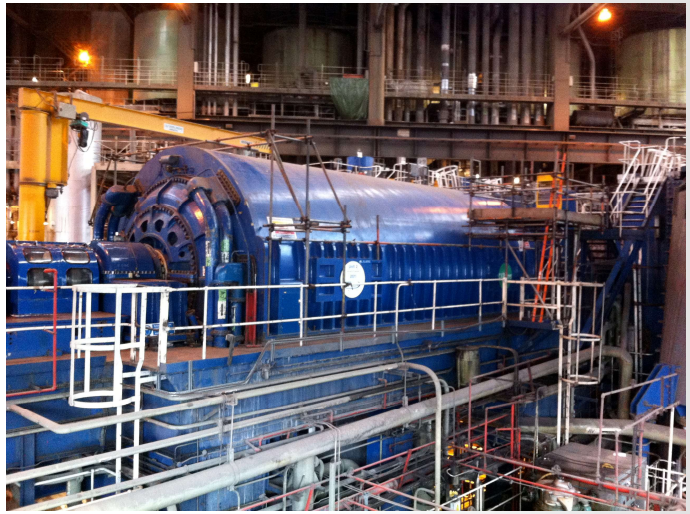



https://abad1dea.tumblr.com/post/182455506350/how-math-can-be-racist-giraffing

PD Data Classification



Partial Discharge

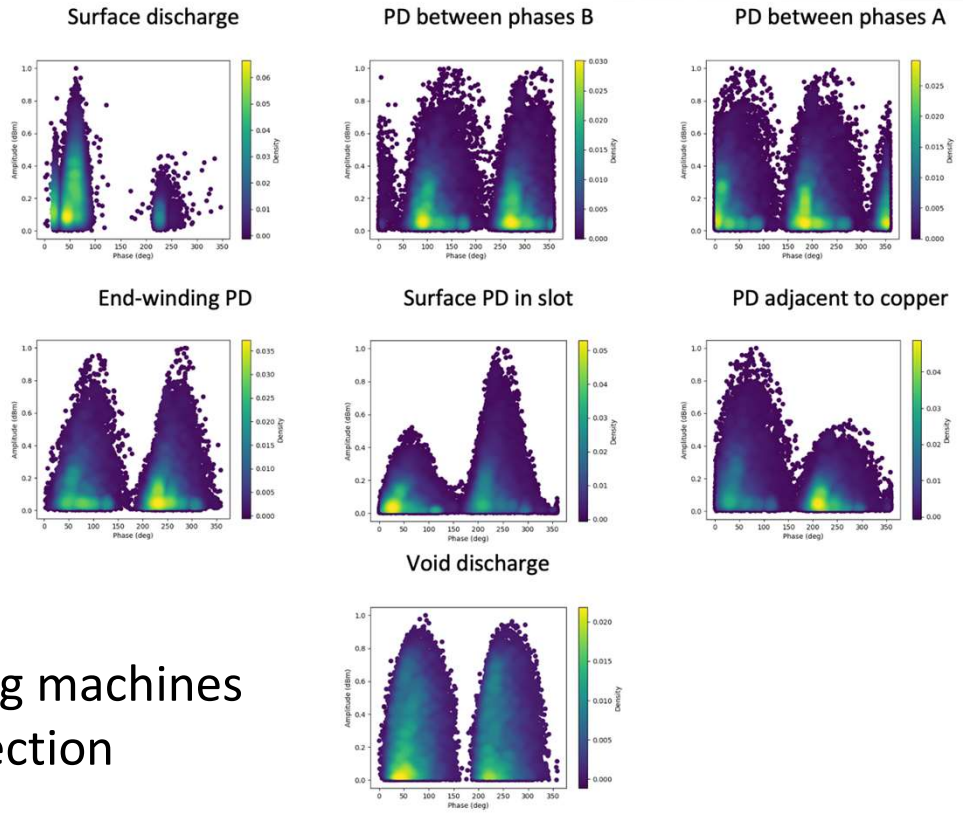


Rotating Machine for Power generation

- Partial Discharge is a symptom of possible incipient failure modes being in operation
- Classification of PD usually requires expert analysis

Phase Resolved Partial Discharge (PRPD) Patterns

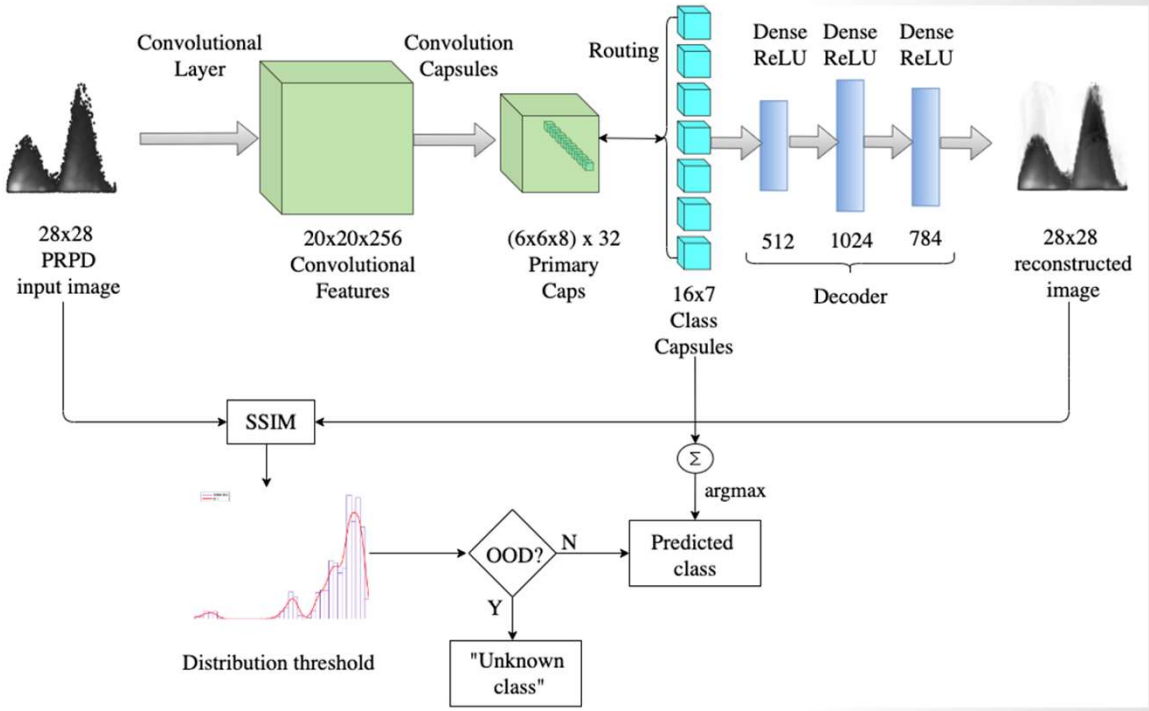
- The Data:
  - Many ‘pulses’ per cycle
  - Records over many cycles
  - Plot size of individual pulses
- Example data shown in dBm
  - Could be in pC
- Patterns are used by ‘experts’ to identify types of PD



- Aims of the work:
  - PD type classification in rotating machines
  - Out-of-Distribution (**OOD**) detection

Method: Classification and OOD

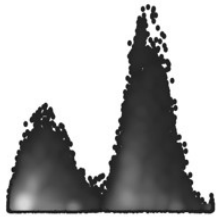
- No point in classifying an image if it is not well represented in the training data set
- The data is 'compressed' and then reconstructed using the 'knowledge' within the AI based on the training data
- Poor reconstruction means the data is not well represented in the training set and may be an anomaly, an outlier, a new 'class' etc



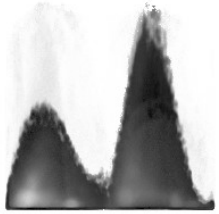
Examples: In and Out of Distribution

In-distribution example

Original Surface PD



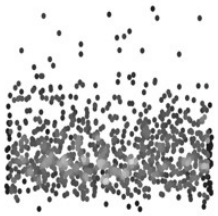
Reconstructed Surface PD



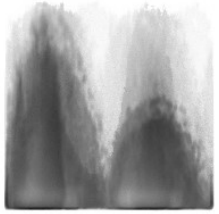
AI trained on cat and dog pictures

OOD example

Original noise



Reconstructed noise



Show it an alligator…

Calculating error for reconstruction

MAE: Mean Absolute Error

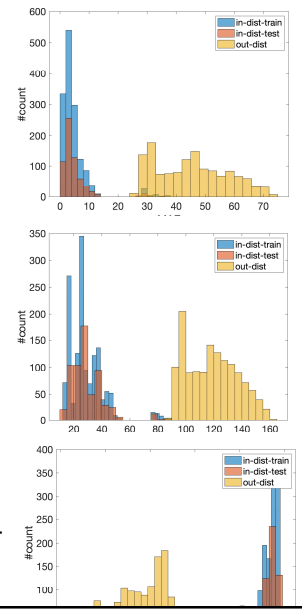
$$MAE = \frac{1}{NL}\sum_{i=0}^{N}\sum_{j=0}^{L}|\boldsymbol{X}(i,j) - \boldsymbol{I}(i,j)|$$

MSE: Mean Squared Error

$$MSE = \frac{1}{NL}\sum_{i=0}^{N}\sum_{j=0}^{L}[\boldsymbol{X}(i,j) - \boldsymbol{I}(i,j)]^2$$

SSIM: Structural Similarity Index Metric

$$SSIM = \frac{(2\mu_{\boldsymbol{X}}\mu_{\boldsymbol{I}} + z_1)(2\sigma_{\boldsymbol{XI}} + z_2)}{(\mu_{\boldsymbol{X}}{}^2 + \mu_{\boldsymbol{I}}{}^2 + z_1)(\sigma_{\boldsymbol{X}}{}^2 + \sigma_{\boldsymbol{I}}{}^2 + z_2)}$$



$\boldsymbol{X}$ input original image

$\boldsymbol{I}$ output reconstructed image

NxL Image dimensions

$z_1 = (0.01S)^2$

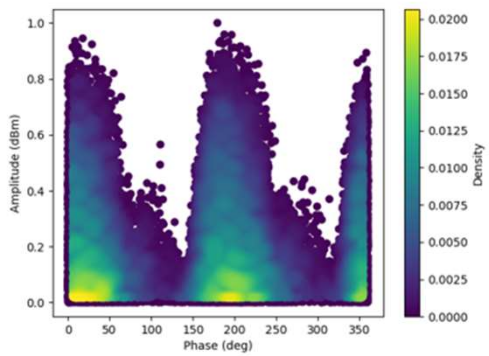
$z_2 = (0.03S)^2$

$S = (2^{\#\text{bits per pixel}} - 1)$

Method	Overall accuracy %	In-distribution accuracy %	OOD accuracy %
CapsNet-MAE	96.61	93.23	100
CapsNet-MSE	96.67	93.34	100
CapsNet-SSIM	**97**	**94**	100

Data from 106 MVA GE Steam Turbine Generator

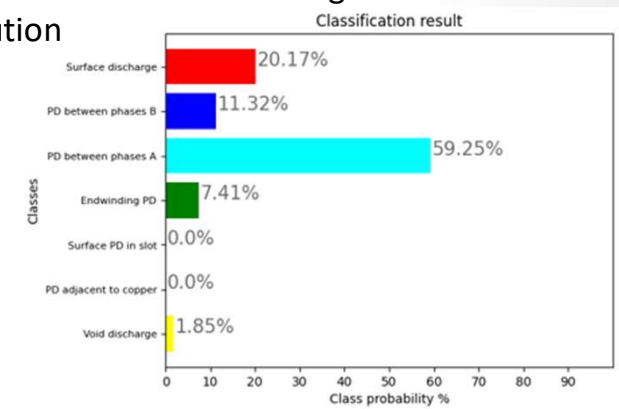
Raw Data

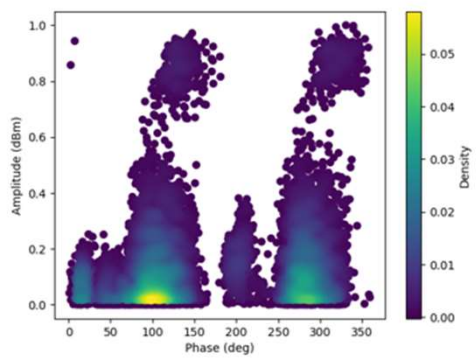


SSIM evaluation of ID or OOD:
In Distribution or Out of Distribution

SSIM 0.94
=> In line with training data

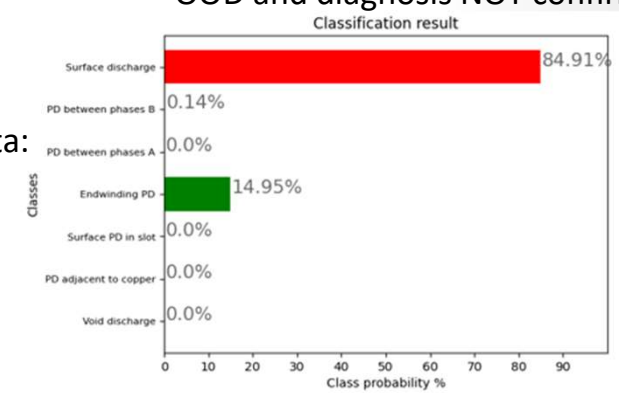
ID and diagnosis confirmed





SSIM 0.38
=> NOT in line with training data:
Out of Distribution

OOD and diagnosis NOT confirmed



Discussion

- OOD detection method shows promise in identifying data which does not have good representation in the original ML data set
- The SSIM metric of the original and the reconstructed PRPD images is a good indicator of OOD
- Note: the OOD data is ‘unseen’ during model’s training
- Successful OOD detection and PD types classification in real-world data – but need to expand the training data set

- Parallel: train a ML system to distinguish between cats and dogs… and the system can also learn to not try to classify an alligator as it doesn’t look at all familiar…

Conclusions

- Hybrid Machine Learning:
  - Make best use of SMEs
  - Use ‘known’ rules/standards/guides
  - Focus the ML on specific problems
  - Iterate

- Out of Distribution Detection
  - Cats, Dogs… Alligators? Giraffes?
  - The ML only performs as well as the training data allows
  - Focus the ML on specific problems

THANK YOU!

Questions? Comments? Feedback?

Tony McGrail

tmcgrail@doble.com

'Always appreciate the opportunity to learn new things.'

Bonus Slide

- **Numbers don't always tell the truth** **(https://qz.com/1427621/companies-are-on-the-hook-if-their-hiring-algorithms-are-biased/)**
- Mark J. Girouard, an employment attorney at Nilan Johnson Lewis, says **one of his clients was vetting a company selling a resume screening tool**, but didn't want to make the decision until they knew what the algorithm was prioritizing in a person's CV.
- After an audit of the algorithm, the **resume screening company found that the algorithm found two factors to be most indicative of job performance: their name was Jared, and whether they played high school lacrosse.** Girouard's client did not use the tool.
- "It's a really great representation of part of the problem with these systems, that **your results are only as good as your training data**," Girouard said. "There was probably a hugely statistically significant correlation between those two data points and performance, but you'd be hard-pressed to argue that those were actually important to performance."