ADVANCING AUTONOMOUS MONITORING & PROGNOSTICS

Novelty Detection and AI-Driven Solutions

Ratna Babu Chinnam, PhD
Professor & Chair
Wayne State University

CAD SEMINAR | FEBRUARY 12, 2025

Industrial & Systems
Engineering Department
Wayne State University
Detroit, MI 48202, USA

ISE.WAYNE.EDU

Background

PROFESSOR & CHAIR
DIRECTOR, AI, BIG DATA & ANALYTICS GROUP
DIRECTOR, INDUSTRYX: CENTER FOR OPERATIONAL EXCELLENCE

Ratna.Chinnam@wayne.edu
Tel: (313) 577-4846

Research Interests:

- **AI & Data Science**: Deep Learning, Data Mining, Operations Research
- **Decision Intelligence**: Decision Analysis
- **Healthcare**: Analytics, Systems Engineering, Decision Support Systems, Access
- **Supply Chain Management**: Supply Chain Modeling & Coordination, Resilience
- **IIoT & Maintenance**: Autonomous Monitoring, Diagnostics, Prognostics
- **Manufacturing**: Industry 4.0, Intelligent Manufacturing, Remanufacturing
- **Product Development**: Product Planning, Lean PD, Design for Sustainability

Scholarship:

- >150 Publications; ~60 PhD Graduates
- Google Scholar: ~6,500 Citations

FEDERAL FUNDING AGENCIES:

INDUSTRY COLLABORATIONS:

Outline

- Motivation
- Condition-Based Maintenance: Current State
- Need for Autonomous Monitoring and Prognostics
- Research & Industrial Case Studies
- Conclusion

Why Study Monitoring, Diagnostics, and Prognostics?

Enable early detection of issues and optimize performance across industries.

- Manufacturing & Industry 4.0
  - Reduce downtime with predictive maintenance
  - Improve quality with AI-driven analytics
  - Optimize efficiency in smart factories
- Automotive & Transportation
  - Prevent costly failures with real-time monitoring
  - Enhance safety in rail and aviation systems
  - Extend the lifespan of EV batteries
- Healthcare & Medical Devices
  - Detect health issues earlier with wearables
  - Improve accuracy in AI-powered diagnostics
  - Personalize treatment with predictive analytics
- Energy & Infrastructure
  - Prevent blackouts with smart grid monitoring
  - Ensure safety with structural health tracking
  - Avoid costly failures in oil & gas pipelines
- Aerospace & Defense
  - Prevent mid-air failures with engine health management
  - Ensure mission success with spacecraft monitoring
  - Detect anomalies before they become threats

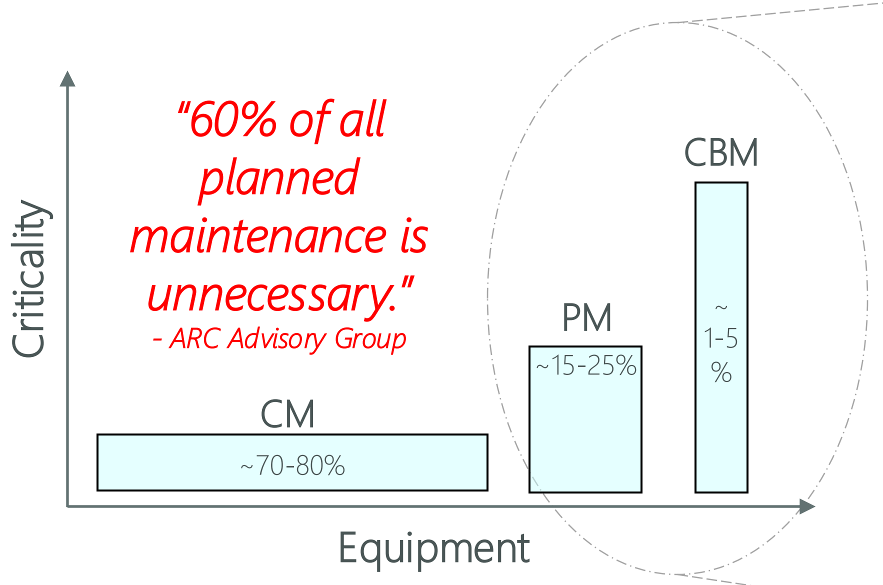
Talk will focus on the Maintenance Domain, but the methods have broader relevance across multiple domains!

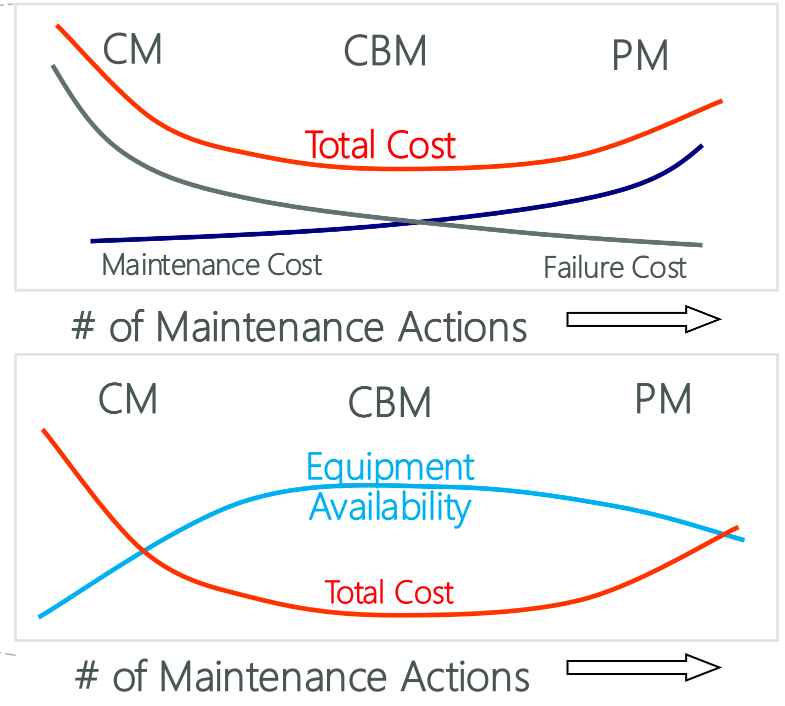
Condition-Based Maintenance

Motivation & Terminology

Background: Maintenance Strategies

- **Corrective Maintenance** (CM): ***Reactive*** action taken only after equipment failure
- **Preventive Maintenance** (PM): ***Routine*** maintenance
- **Condition-Based Maintenance** (CBM): ***Maintenance as needed*** based on real-time equipment conditions, preventing unnecessary interventions
  - Involves monitoring of asset's sensors and signal analysis

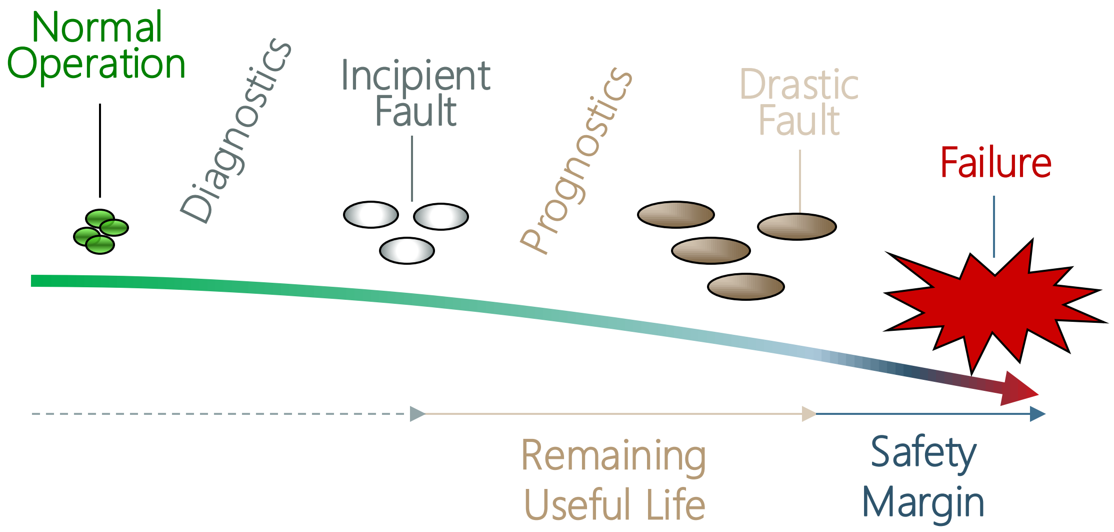




ARC Advisory Group is a research and advisory firm for industry and infrastructure: http://www.arcweb.com

Motivation for CBM Research

- $1 trillion/year is spent replacing well functional equipment
  - Due to lack of reliable methods to predict remaining useful life (McLean & Wolfe, 2002).
- Effective CBM technology could save $35B annually in US alone (Lee, 2003).
- Unplanned downtime costs exceed $100k/hour
  - Makes accurate failure prediction more critical than ever (PM&APM Report, 2023).



TERMINOLOGY:

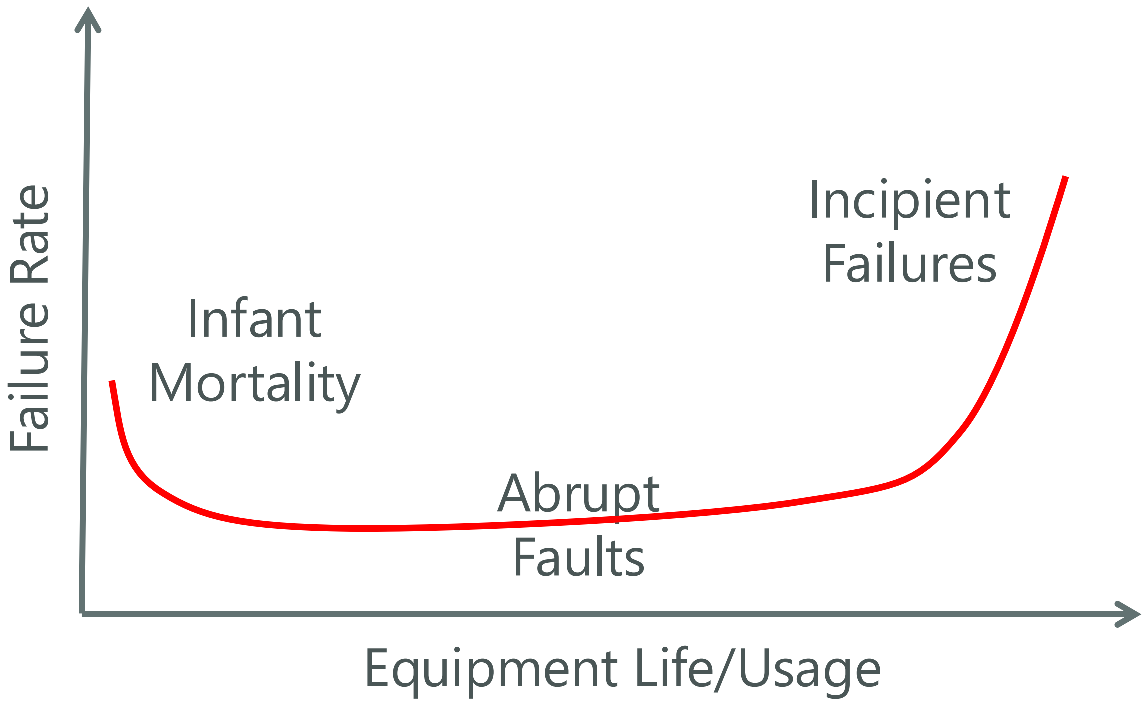
MONITORING: ability to track asset (sensor signals) for anomalous behavior

DIAGNOSTICS: ability to detect and classify fault conditions

PROGNOSTICS: ability to predict progression of fault condition to failure

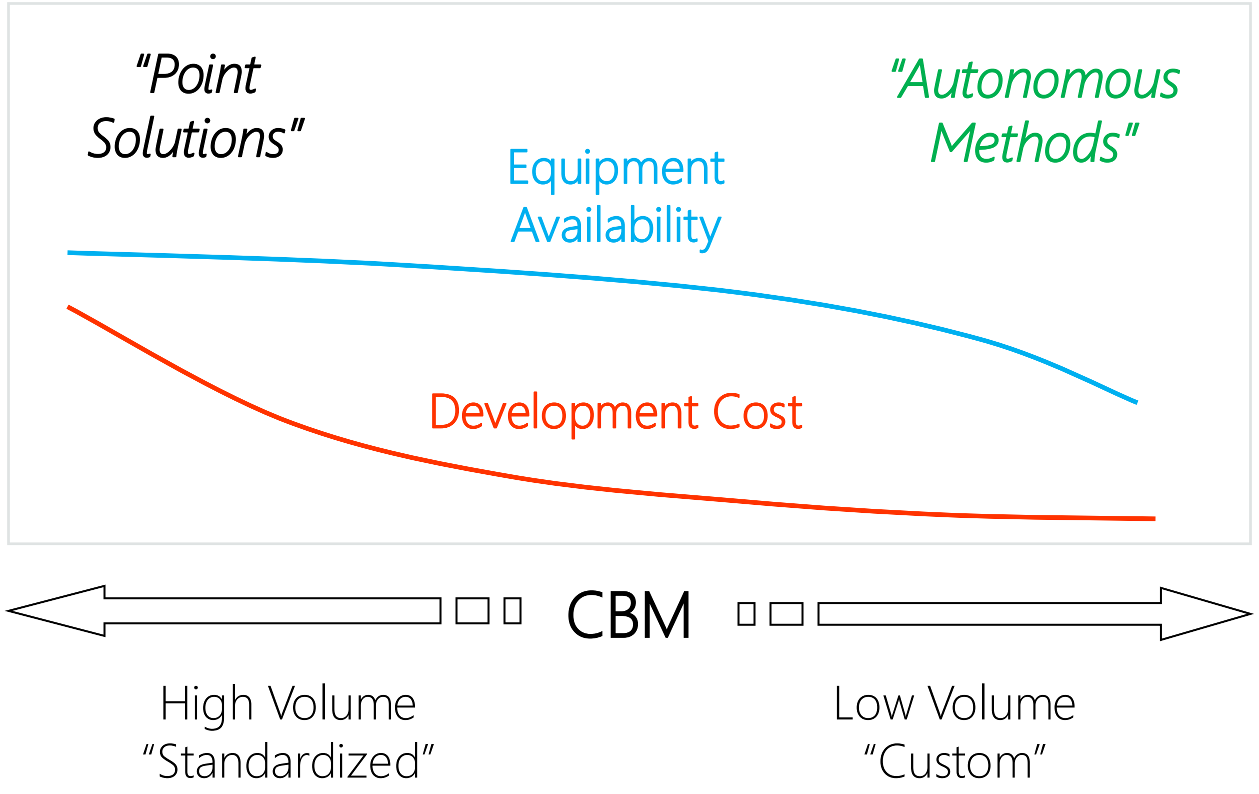
Failure Types & Monitoring Strategy

- Infant Mortality
  - Failure due to manufacturing problems
  - Strategy: End-of-line testing and burn-in strategies are effective.
- Abrupt Faults
  - Occur in a short time
  - Difficult to track development
  - Strategy: Monitoring and novelty detection techniques.
- Incipient Failures
  - Occur slowly due to "wear and tear"
  - Possible to track development
  - Strategy: Health-state estimation techniques and prognostics.



CBM: Point-Solutions vs. Autonomous Approaches

- Conventional "*point solution*" *methods* rely on extensive domain knowledge, characterization, making them time-consuming and costly.
- As industries evolve rapidly, there is a growing need for *adaptable solutions*.
- We need "*generic*" *autonomous methods* that are rapidly configurable, learn well from data, and can handle a wide variety of equipment/components.



CBM: Core Elements

Asset of Interest

(Sensors)

OPTIONAL

Sensor Signals → **Feature Extraction** (OPTIONAL) → **Monitoring** → **Diagnostics** → **Prognostics** → **Recommended Maintenance Action**

- Sensor Signals
  - Different Types
  - Variable Fidelity
  - Conditioning & Fusion
- Feature Extraction
  - Feature Engineering
  - Feature Selection
  - Dimensionality Reduction
- Monitoring
  - Novelty Detection
- Diagnostics
  - Fault Classification
  - Fault Condition
  - Fault Isolation
- Prognostics
  - Remaining Useful Life
  - Next Probable States
  - Next Probable Faults

Advancing Autonomous Monitoring, Diagnostics & Prognostics

Research Goals

- Develop Autonomous Approaches
  - Create adaptable, generic, and autonomous methods for monitoring and prognostics to facilitate CBM.
- Overcome Data Quality Challenges
  - Address issues related to poor data quality, low fidelity, data sparsity, and the impact of external environmental factors.
- Create Flexible Frameworks
  - Design end-to-end frameworks capable of managing variability and complexity across diverse systems and application settings.
- Ensure Real-World Applicability
  - Validate solutions through rigorous testing across multiple case studies to ensure practical relevance and effectiveness.

Feature Engineering

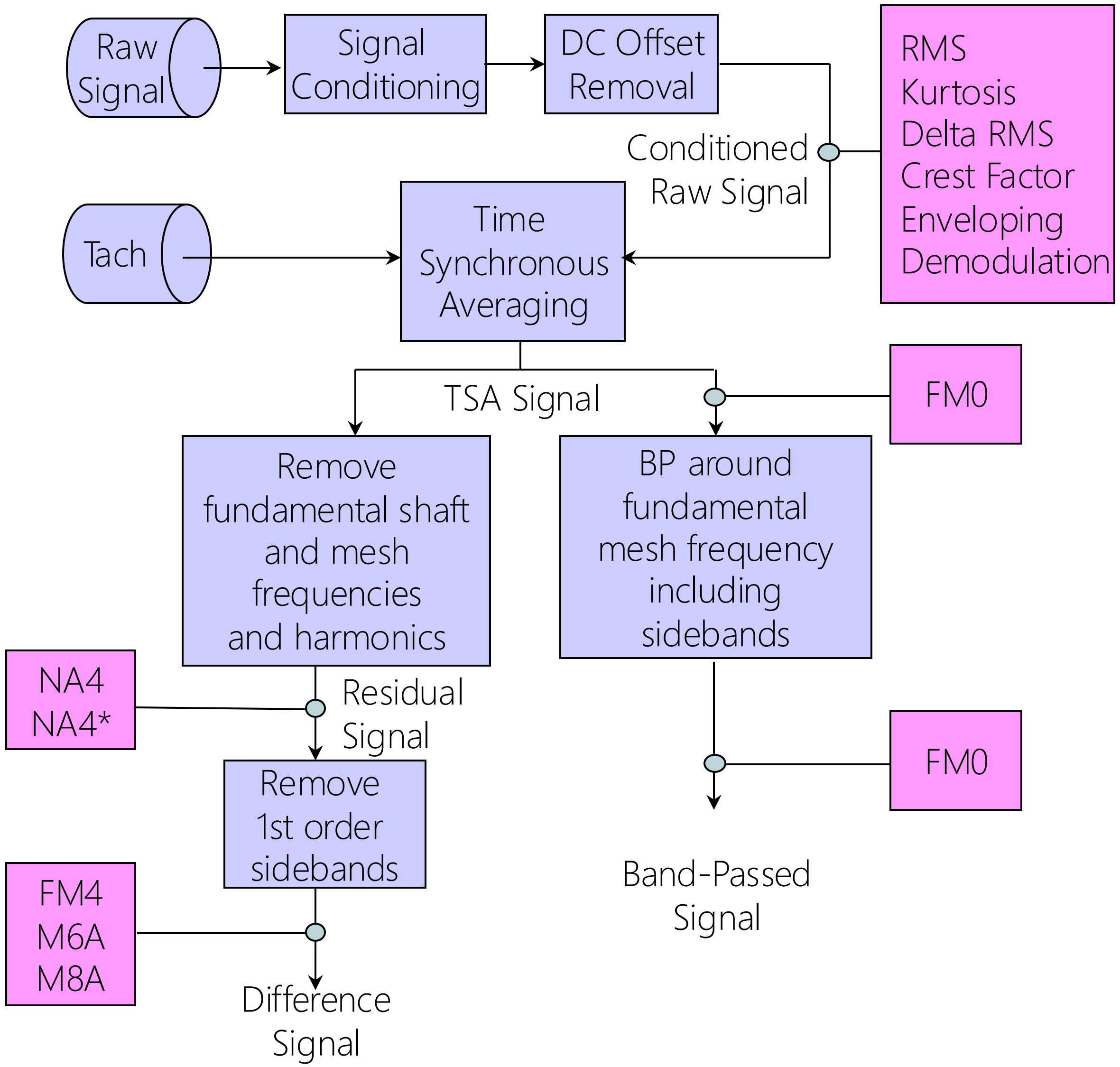
Feature Engineering

- **Definition**: Process of selecting, transforming, and creating features (inputs) to improve the performance of models.
  - Can be broadly grouped into time-, frequency-, and mixed-domain methods.
- **Challenges**:
  - Requires domain expertise to identify relevant features (e.g., vibration signals).
  - Time-consuming and labor-intensive.
  - Complex and may lead to missed insights due to manual limitations.

ILLUSTRATIVE EXAMPLE: GEARBOXES

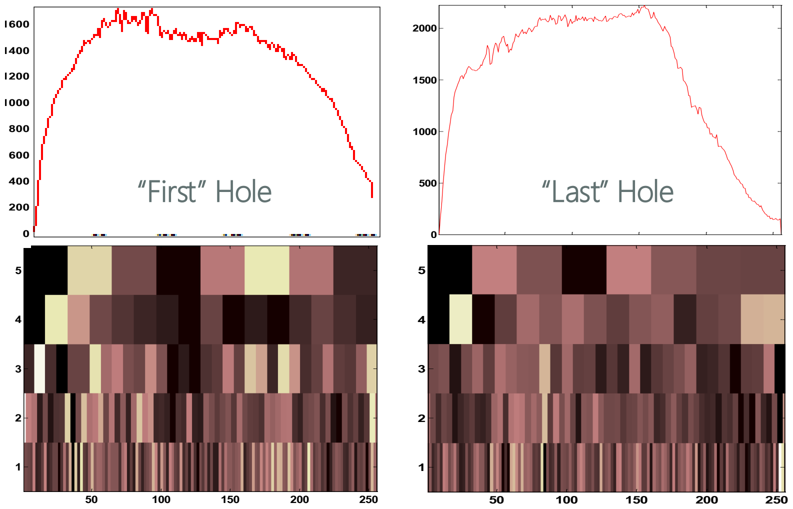
Process Flow for CBM Feature Extraction Methods

(Source: Lebold et al., 2000)



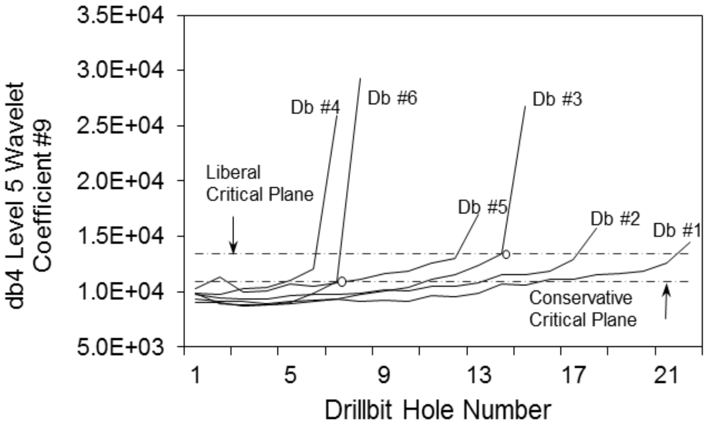
Feature Extraction for Monitoring Cutting Tools

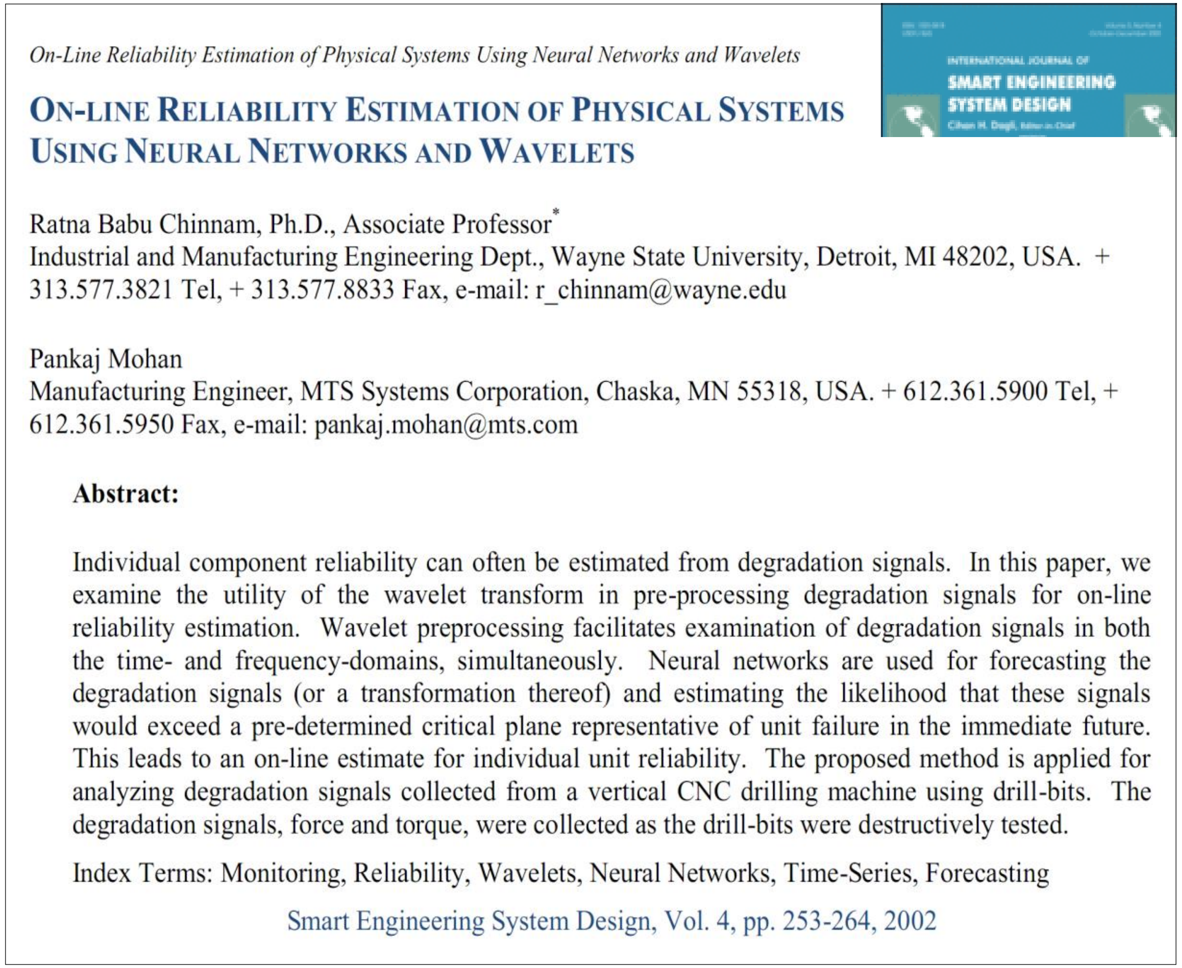
Thrust Force Signals from Drilling



DB4 Wavelet Coefficients

Single DB4-Level 5 Wavelet Coefficient Adequate for Tool Monitoring & Optimal Replacement





On-Line Reliability Estimation of Physical Systems Using Neural Networks and Wavelets

INTERNATIONAL JOURNAL OF SMART ENGINEERING SYSTEM DESIGN

ON-LINE RELIABILITY ESTIMATION OF PHYSICAL SYSTEMS USING NEURAL NETWORKS AND WAVELETS

Ratna Babu Chinnam, Ph.D., Associate Professor*
Industrial and Manufacturing Engineering Dept., Wayne State University, Detroit, MI 48202, USA. + 313.577.3821 Tel, + 313.577.8833 Fax, e-mail: r_chinnam@wayne.edu

Pankaj Mohan
Manufacturing Engineer, MTS Systems Corporation, Chaska, MN 55318, USA. + 612.361.5900 Tel, + 612.361.5950 Fax, e-mail: pankaj.mohan@mts.com

Abstract:

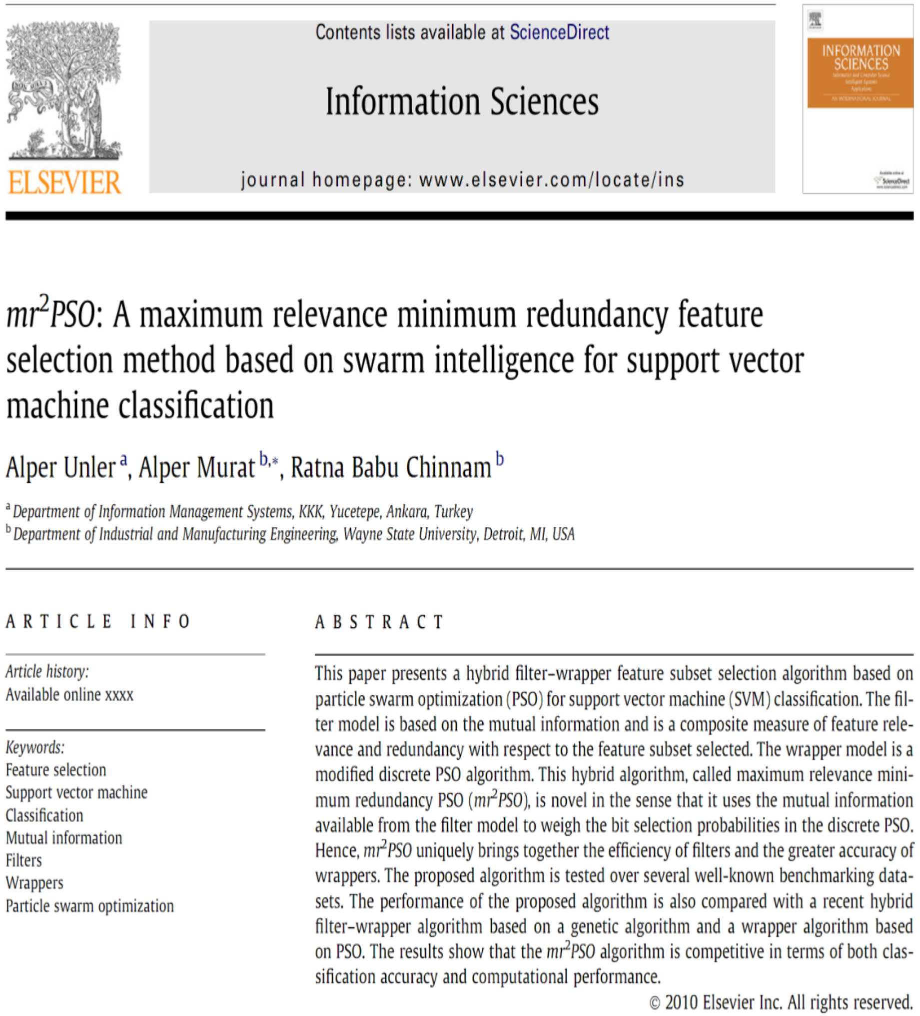
Individual component reliability can often be estimated from degradation signals. In this paper, we examine the utility of the wavelet transform in pre-processing degradation signals for on-line reliability estimation. Wavelet preprocessing facilitates examination of degradation signals in both the time- and frequency-domains, simultaneously. Neural networks are used for forecasting the degradation signals (or a transformation thereof) and estimating the likelihood that these signals would exceed a pre-determined critical plane representative of unit failure in the immediate future. This leads to an on-line estimate for individual unit reliability. The proposed method is applied for analyzing degradation signals collected from a vertical CNC drilling machine using drill-bits. The degradation signals, force and torque, were collected as the drill-bits were destructively tested.

Index Terms: Monitoring, Reliability, Wavelets, Neural Networks, Time-Series, Forecasting

Smart Engineering System Design, Vol. 4, pp. 253-264, 2002

Feature Selection Methods

- **Motivation: Hughes Phenomenon**
  - Definition: As the number of features increases, model performance degrades beyond a certain threshold, given a fixed sample size.
- **Key Techniques Involve:**
  - **Search Techniques**: Finding the best subset of features.
  - **Evaluation Measures**: Assessing quality of each subset.
- **Types of Feature Selection:**
  - **Filters**: Use a proxy measure (e.g., correlation, mutual information) rather than model error rate to rank features.
    - Advantages: Fast, scalable, independent of predictive models.
  - **Wrappers**: Train a predictive model on different feature subsets and select the one with the highest performance.
    - Advantages: Typically yields better performance than filters but computationally expensive.
  - **Embedded Methods**: Feature selection occurs as part of the model training process.
    - Examples: Lasso (L1 regularization), Decision Trees



Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

ELSEVIER

mr^2*PSO*: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification

Alper Unler [a], Alper Murat [b,*], Ratna Babu Chinnam [b]

[a] Department of Information Management Systems, KKK, Yucetepe, Ankara, Turkey
[b] Department of Industrial and Manufacturing Engineering, Wayne State University, Detroit, MI, USA

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Feature selection
Support vector machine
Classification
Mutual information
Filters
Wrappers
Particle swarm optimization

ABSTRACT

This paper presents a hybrid filter–wrapper feature subset selection algorithm based on particle swarm optimization (PSO) for support vector machine (SVM) classification. The filter model is based on the mutual information and is a composite measure of feature relevance and redundancy with respect to the feature subset selected. The wrapper model is a modified discrete PSO algorithm. This hybrid algorithm, called maximum relevance minimum redundancy PSO (*mr*^{2}*PSO*), is novel in the sense that it uses the mutual information available from the filter model to weigh the bit selection probabilities in the discrete PSO. Hence, *mr*^{2}*PSO* uniquely brings together the efficiency of filters and the greater accuracy of wrappers. The proposed algorithm is tested over several well-known benchmarking datasets. The performance of the proposed algorithm is also compared with a recent hybrid filter–wrapper algorithm based on a genetic algorithm and a wrapper algorithm based on PSO. The results show that the *mr*^{2}*PSO* algorithm is competitive in terms of both classification accuracy and computational performance.

Proposed a "hybrid" framework that brings the efficiency advantage of filters with the accuracy performance of wrappers!
>375 Citations on Google Scholar

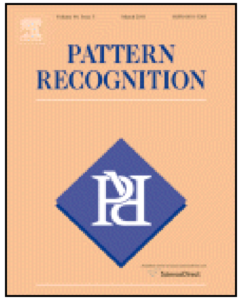
End-to-End AI Pipelines for Feature Engineering

- Automated Feature Extraction:
  - AI models, particularly deep learning models, can potentially extract features from raw data, reducing reliance on manual feature engineering.
- End-to-End Learning:
  - AI systems learn directly from raw sensor data, combining feature extraction and prediction into a single process.
- Advantages:
  - Scalability: Can be applied across multiple systems and datasets without manual intervention.
  - Efficiency: Accelerates deployment by reducing the need for human-in-the-loop feature selection.
  - Improved Performance: Learns hidden, complex patterns not easily captured by traditional feature engineering methods.

We are currently developing deep learning methods to directly generate optimal maintenance plans for the energy industry, such as wind farms, using raw sensor data from assets.

Novelty Detection

General Support Vector Representation Machines



Pattern Recognition 41 (2008) 3021 – 3034

General support vector representation machine for one-class classification of non-stationary classes

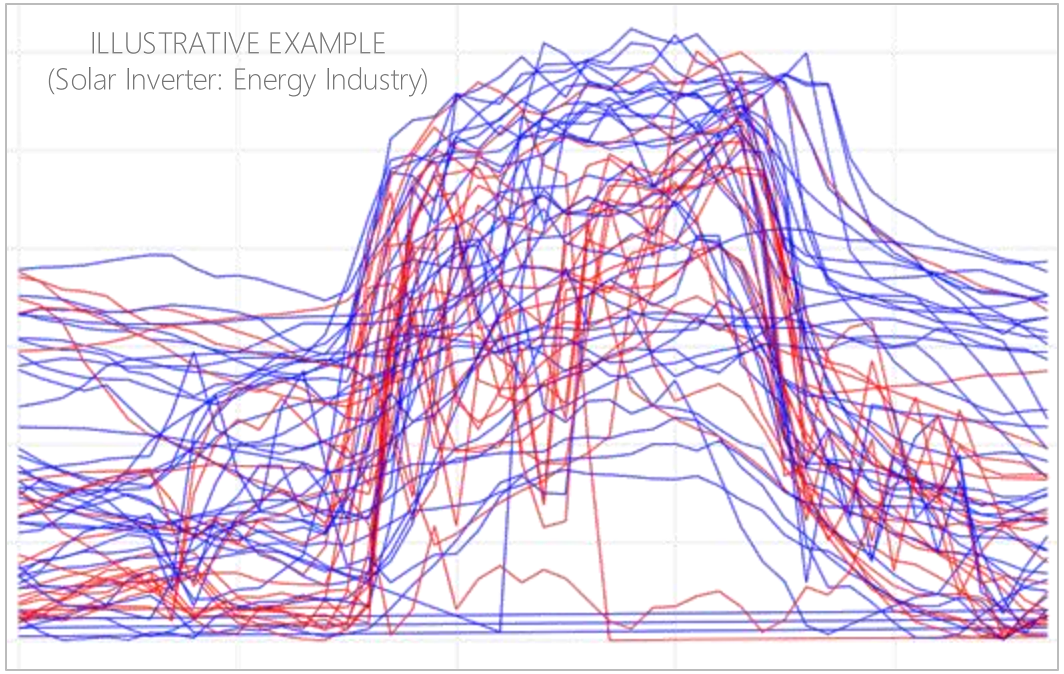
Fatih Camci[a], Ratna Babu Chinnam[b,*]

[a]Department of Computer Engineering, Fatih University, Buyukcekmece, Istanbul, 34500 Turkey
[b]Department of Industrial and Manufacturing Engineering, Wayne State University, 4815 Fourth Street, Detroit, MI 48202, USA

Introduction

Monitoring and Novelty Detection

- **Novelty Detection**: "The task of identifying when test data differ in some respect from the normal data available during training." (Pimentel et al. 2014)
- It is typically approached as a **one-class classification** problem. (Moya et al. 1993)
- **Confounding factors,** such as environmental variables, can obscure real novelties.
- **Each novelty detection method has its own strengths and limitations.**
- **Ensembles methods can enhance detection performance** by combining the advantages of multiple approaches.



How to distinguish anomalous condition states from normal states?

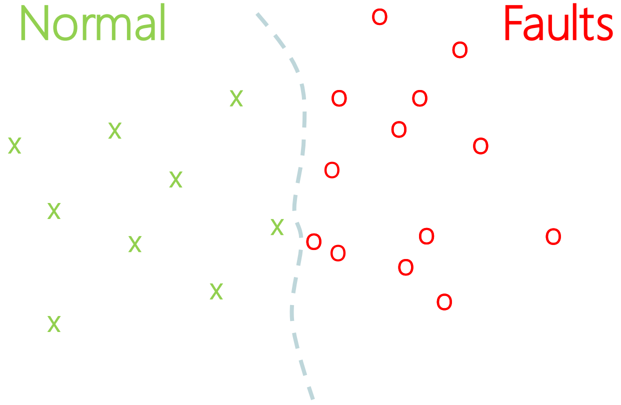
Pattern Recognition vs. Novelty Detection

Pattern Recognition Approach:

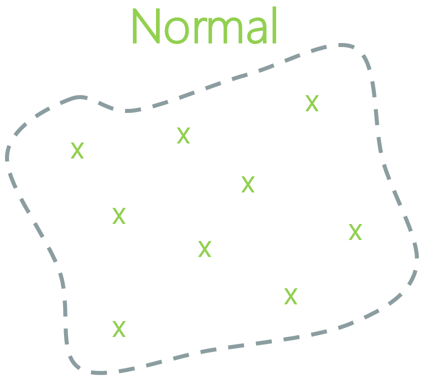
- Relies on labeled examples from all fault classes, including both normal and abnormal conditions.
- Challenges:
  - Gathering examples for all potential fault classes is difficult.
  - Time-consuming to generate examples of rare/unknown failure modes.

Novelty Detection (One-Class Classification) Approach:

- **Focuses on learning and modeling only the "normal" operating behavior** of the system. When system deviates from normal behavior, it is flagged a **potential anomaly** or fault.
- Advantages:
  - Does not require prior knowledge of all failure types.
  - Ideal for monitoring systems where anomalies are rare but critical.



2-Class Classification Example



One-class Classification

One-Class Classification for Anomaly Detection

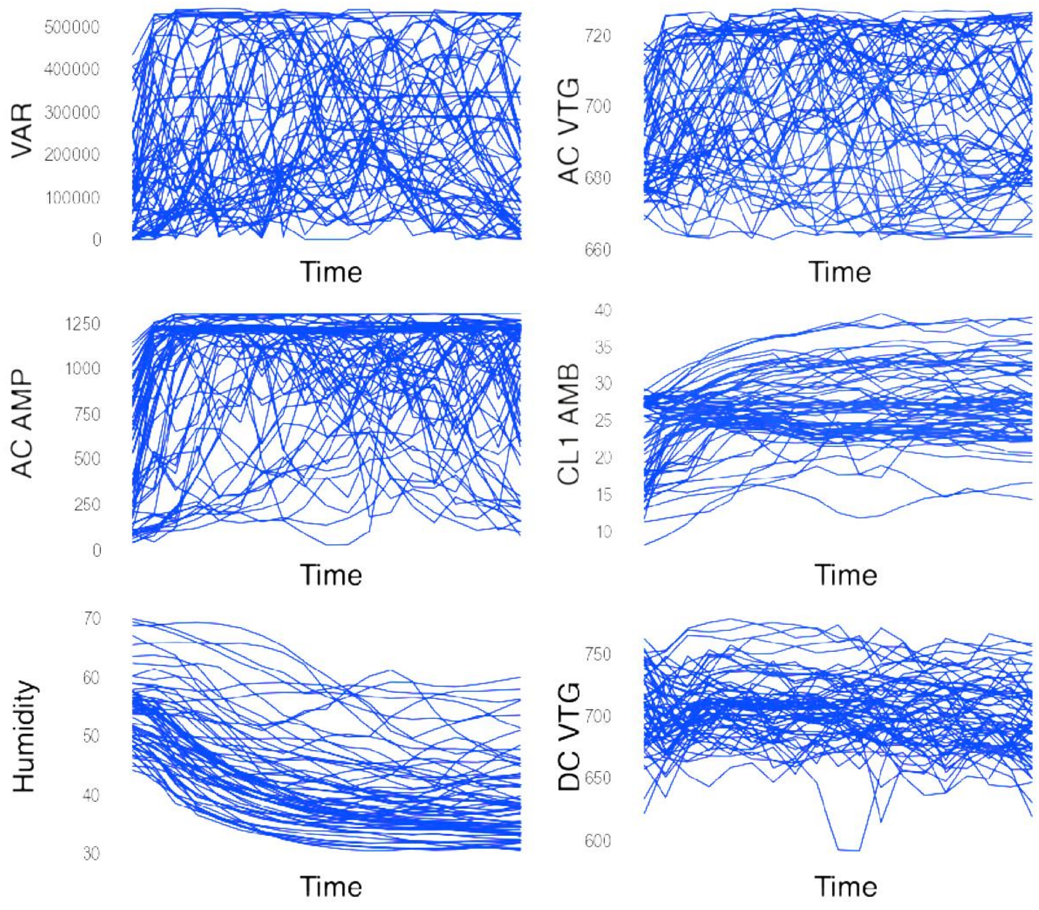
- Problem Definition:
  Given a dataset $X = \{x_1, x_2, \ldots, x_n\}$, containing only normal data points $x_i \in \mathbb{R}^d$, the goal is to learn a function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ to detect anomalies based on deviations from normal behavior, without assuming any specific data distribution.
- Training:
  Learn a decision function f that models the normal class, and define a decision threshold θ to classify data as normal or anomalous.
- Detection:
  For a new data point x_{test}:

$$\hat{y} = \begin{cases} 1 & \text{if } f(x_{test}) \geq \theta \quad \text{(anomalous)} \\ 0 & \text{if } f(x_{test}) < \theta \quad \text{(normal)} \end{cases}$$

- Objective:
  Minimize false positives and maximize the detection of true anomalies, learning a boundary that generalizes well to unseen data.

Need Meaningful Features

- **Extracting meaningful information** from raw data is crucial for effective novelty detection.
- Novelties **can exhibit dynamic behavior** and evolve over time.
- In temporal data sets, **sequential observations are more telling** than point observations.

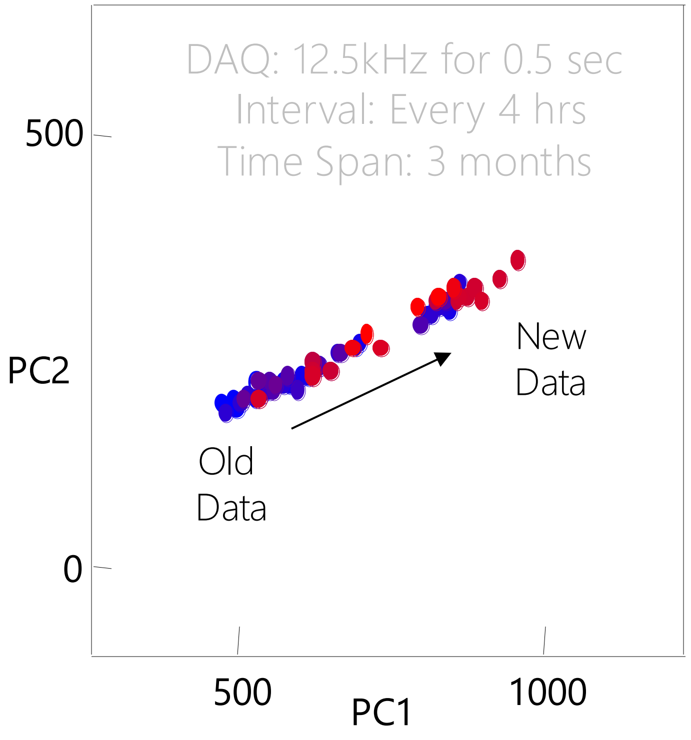


Daily Signals from a Solar Inverter

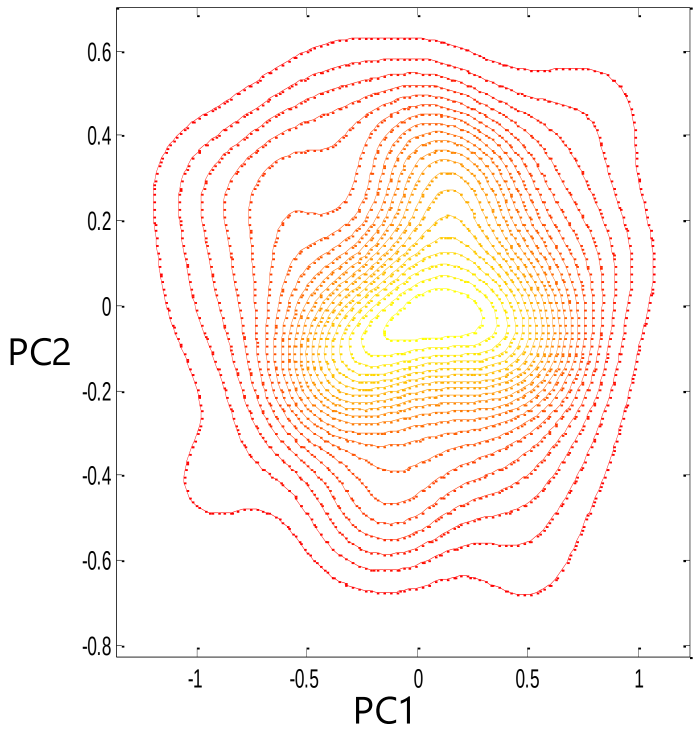
(each line represents observations of a day)

Difficulties with One-Class Classification

- Inability to **handle non-stationary processes**
- **Unrealistic assumptions**
  - Example: Data density, Independence
- **Inability to exploit any available “limited” data from fault classes**



2 PCs of Principal Spectral Frequencies of Vibration Sensor Data Collected from a Pump



Kernel Density Contours of 2 PCs Temporal Domain

General Support Vector Representation Machine

- GSVRM: Minimize volume hyper-sphere containing “normal” data

*“**Primal**” Formulation:*

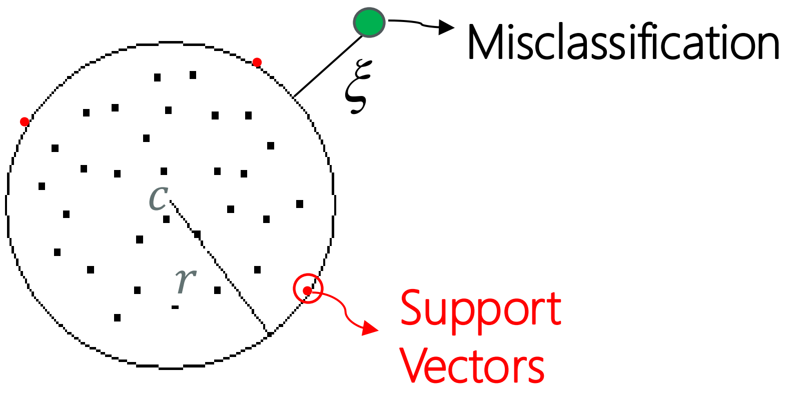
Min $\quad r^2 + C\sum_i \xi_i \quad$ s.t. $\quad \|x_i - c\|^2 \le r^2 + C\sum_i \xi_i$

$\xi_i \ge 0$

*“**Dual**” Formulation:*

Max $\quad \sum_i \alpha_i (x_i \cdot x_i) - \sum_{i,j} \alpha_i \alpha_j (x_i \cdot x_j)$

s.t. $\quad 0 \le \alpha_i \le C \qquad \sum_i \alpha_i = 1$

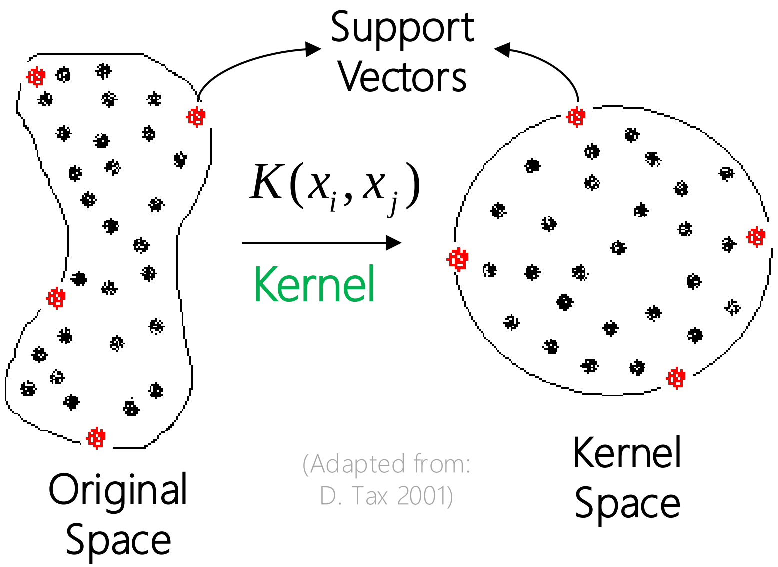


- Non-Spherical Data: Rely on “Kernels”

Mapping: $\quad x_i \cdot x_j \longrightarrow K(x_i, x_j)$

Max $\quad \sum_i \alpha_i K(x_i, x_i) - \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j)$

s.t. $\quad 0 \le \alpha_i \le C \qquad \sum_i \alpha_i = 1$



(Adapted from: D. Tax 2001)

Non-stationary Processes: Adaptive-GSVRM

- Primal Formulation:

Min $\quad r^2 + C\sum_i \boxed{\omega_i}\, \xi_i$

s.t. $\quad \|x_i - c\|^2 \le r^2 + C\sum_i \omega_i \xi_i, \forall i$

$$\boxed{\omega_{t_c - i} = (1-\lambda)^{t_c - i}}$$

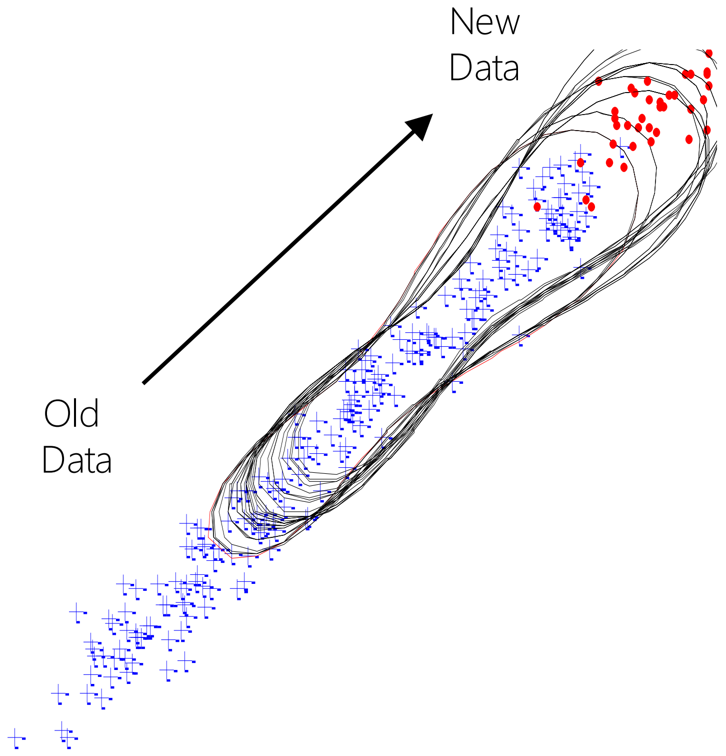
Still remains a quadratic formulation!

- Dual Lagrangian Formulation:

Max $\quad \sum_i \alpha_i K(x_i, x_i) - \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j)$

s.t. $\quad 0 \le \alpha_i \le C\omega_i \qquad 0 < \lambda < 1$

$\sum_i \alpha_i = 1 \qquad \xi_i \ge 0$



Results from Benchmarking Datasets

- Novelty Detection Accuracy Results
  - Average of Type I and Type II errors
- Test Sets:
  - Stationary Processes
    - Normal, log-normal, and exponential distributions
    - Smith dataset [Smith, 1994]
  - Non-stationary Processes
    - Viscosity dataset
    - Papermaking dataset
- Adaptive-GSVRM
  - Comes close to the performance of binary classification ML methods (e.g., SVM, MLP, RBF) with full access to fault data.

Stationary Processes

	Normal	Lognormal	Exponential
GSVRM3	94.9%	90.4%	88.1%
GSVRM1	94.7%	90.0%	88.1%

Smith Data	
SVM	**96.5%**
GSVRM2	**93.2%**
GSVRM3	93.2%
GSVRM1	89.8%
Shewhart Chart	86.5%
MLP	86.0%

Non-Stationary Processes

Viscosity	
SVM	**97.8%**
GSVRM2	**96.5%**
GSVRM3	92.8%
GSVRM1	91.0%
RBF	87.7%

Papermaking	
SVM	**97.5%**
GSVRM2	**95.5%**
RBF	95.0%
GSVRM1	94.5%
GSVRM3	89.8%

GSVRM1: Training with "normal" data alone
GSVRM2: Training with normal and limited failure data
GSVRM3: Training with only 25 normal and 10 abnormal samples

Asset “Health-State” Estimation

Unsupervised Learning with Hidden Markov Models

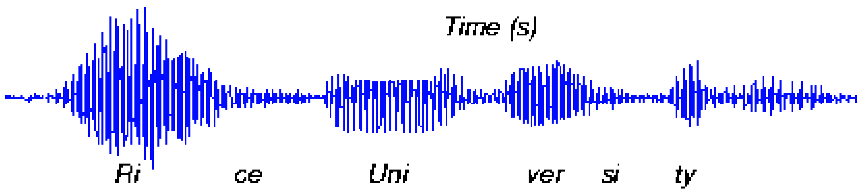
IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING, VOL. 7, NO. 3, JULY 2010 581

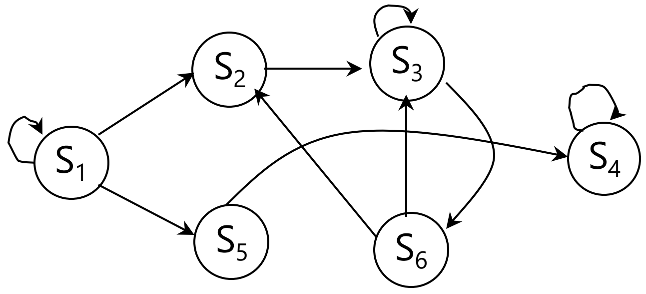
Health-State Estimation and Prognostics in Machining Processes

Fatih Camci and Ratna Babu Chinnam

Hidden Markov Models (HMM)

- **Definition**: Doubly embedded stochastic process with hidden states
  - Underlying Markov process (hidden state sequence) plus stochastic emissions
- **Motivation**:
  - Widely successful in speech recognition (SR) applications
  - Faulty diagnostics has a lot in common with SR





State Transitions

- **HMM Structure**:
  - Initial state distribution: $\pi(i) = P(X_1 = i)$
  - State transition matrix: $A(i,j) = P(X_t = j | X_{t-1} = i)$
  - Observation model: $B = P(O_t | X_t)$

X_t: hidden state at time t

O_t: observation at time t

- **HMM Attributes**:
  - Empirical parametric models that can learn from data
  - Rich mathematical structure and interpretability

Training and Learning HMMs

- Learning Task:
  - Adjust model parameters to maximize likelihood given observation sequences

- Limitations of Standard HMMs:
  - Computationally inefficient
  - Lack of structural flexibility

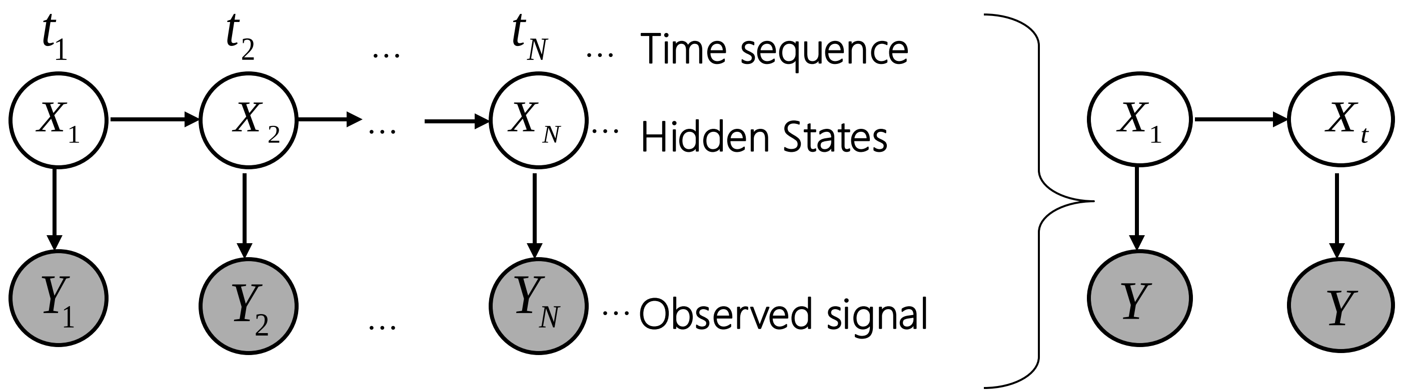
- Dynamic Bayesian Networks (DBNs) as an Alternative:
  - Factored Representation: Uses fewer parameters by decomposing state variables into smaller, manageable components.
  - Structural Flexibility: DBNs allow for more flexible modeling of complex systems with dynamic dependencies, overcoming some of the limitations of HMMs.

Dynamic Bayesian Network (DBN) Representation

- **Temporal Modeling**: DBNs effectively capture how variables evolve over time by modeling dynamic dependencies across time slices.
- **Structure**: Consists of two networks
  - **"Prior" Network**: Encodes prior probabilities for the initial time slice
  - **"Transition" network**: Defines conditional transition probabilities for subsequent time slices
- **Expanding Range of Algorithms**
- **Special cases of DBNs**: Hidden Markov Models, Kalman Filters, …

HMM as DBN

"Roll"



3 Nodes

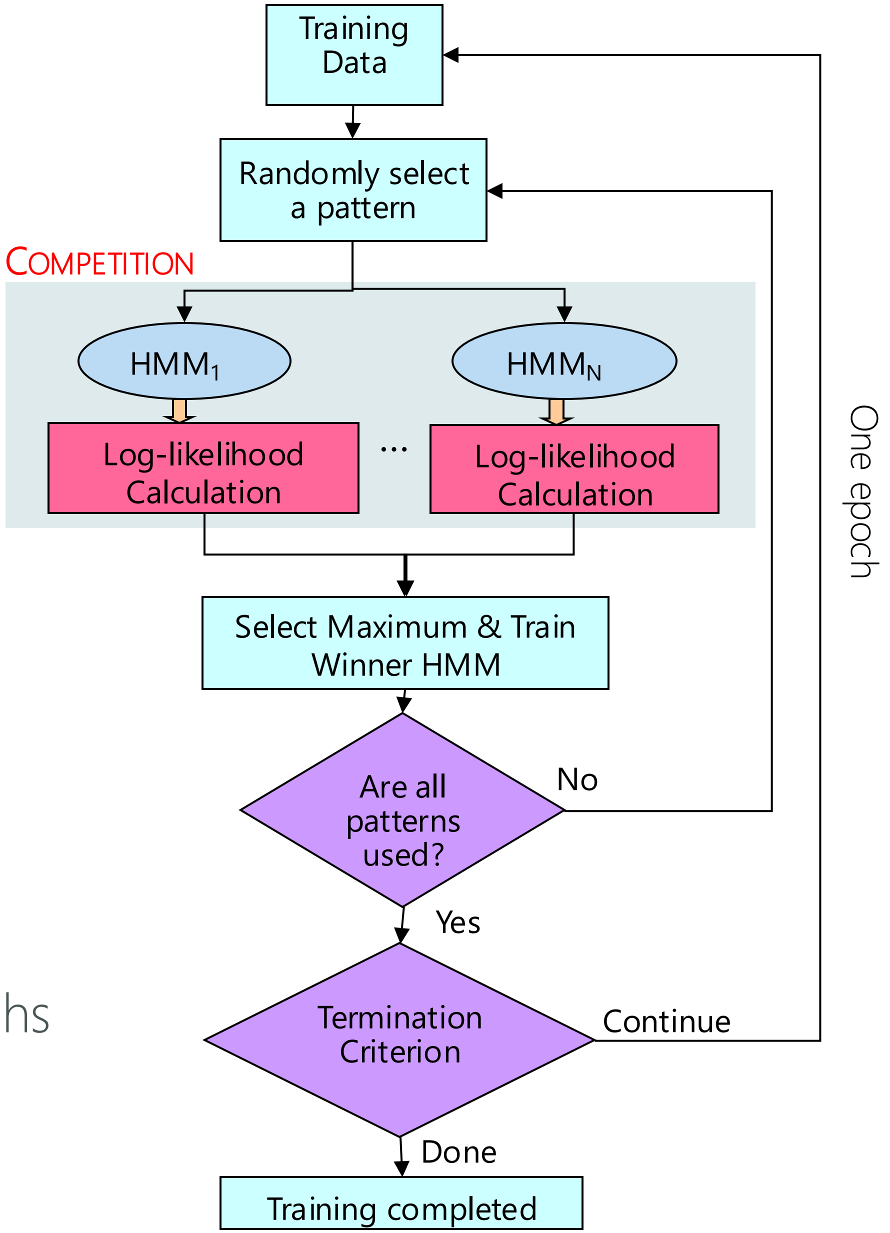
	Parents	
X_1	N/A	
X_t	X_{t-1}	$t = 2, 3 \ldots.$
Y_t	X_t	$t = 1, 2, \ldots.$

$$P\left(Y_t = y_t \mid X_t = i\right) = N\left(\mu_i, \Sigma_i\right)$$

HMMs for Modeling Health-States

"Competitive" Learning

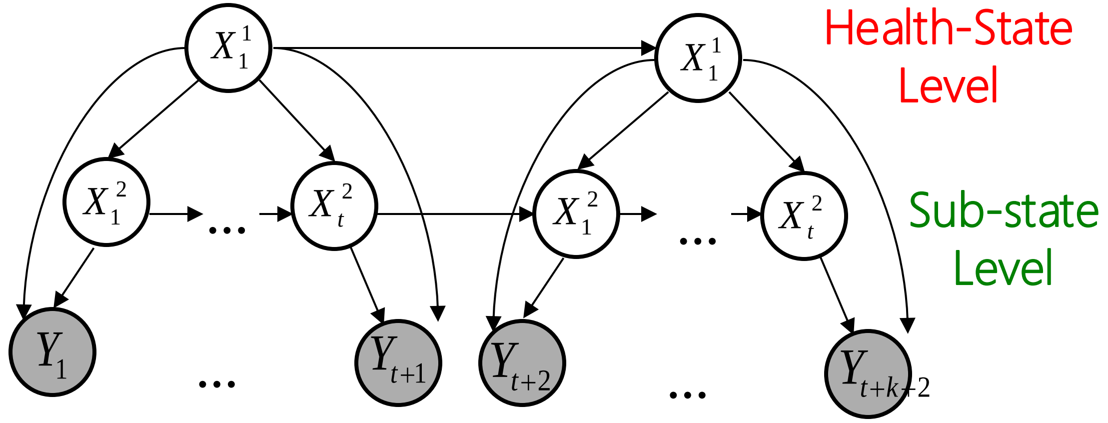
- **Difficulty**: Unlabeled Data
- **Solution**: Model-Based Clustering
- **Competitive Learning**
  - HMM with highest likelihood wins the competition
  - Winner HMM is trained with the data
- **Termination Criteria**
  - Error minimization
    - Training until no reverse jump exists
    - May not be applicable to very noisy data
  - Convergence
    - Training until no change occurs in two consecutive epochs
    - Parameters: learning rate, reduction rate

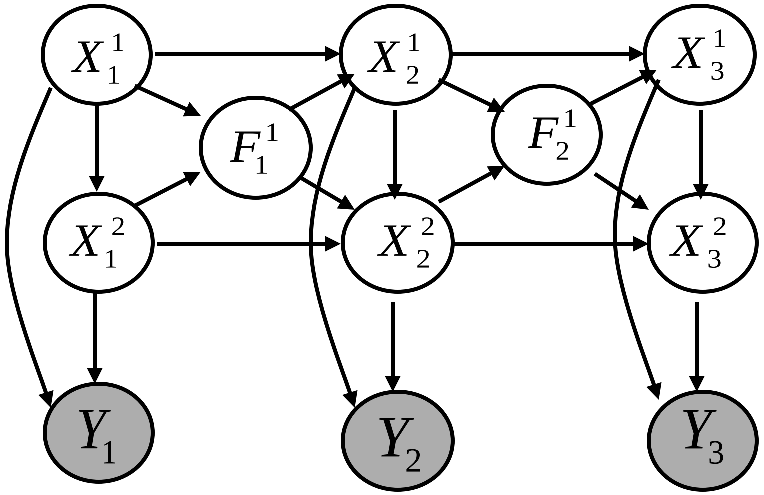


Hierarchical Hidden Markov Models

One-Shot Learning

- **Assumption**:
  - Sensor signal(s) depict health-states
  - Health-states consist of sub-states
- **Difficulty**: Initialization of H-HMM structure





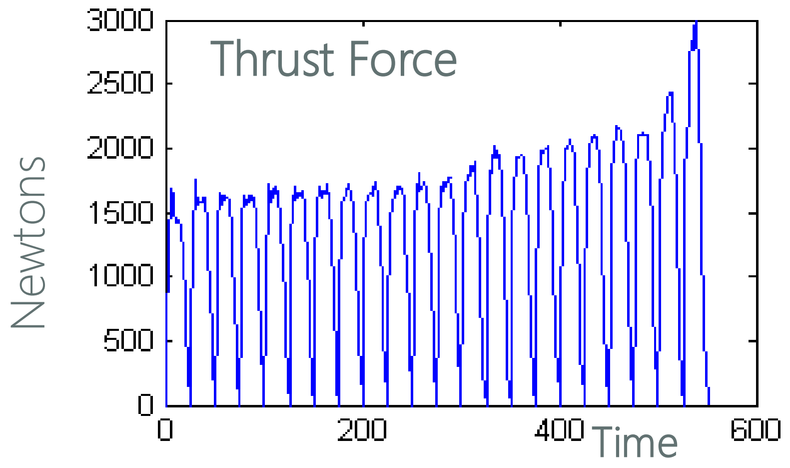
6 Nodes

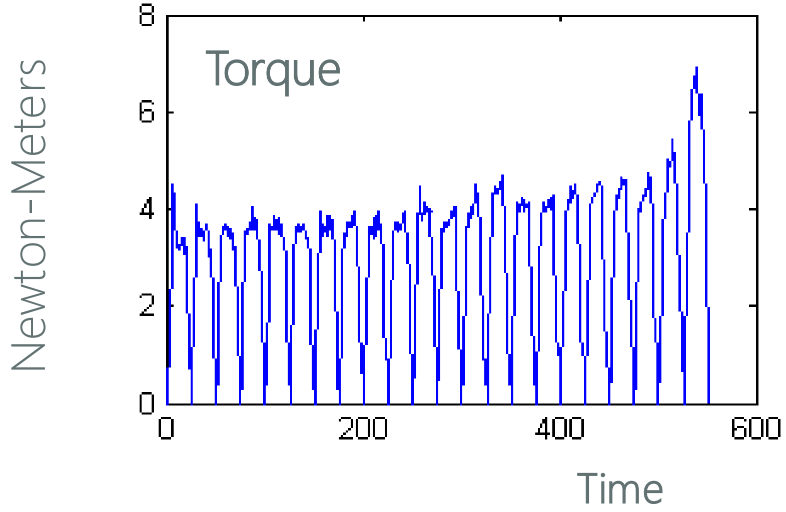
	Parents		Parents	
X_1^1	N/A	F_t^1	$X_{t-1}^1 \quad X_{t-1}^2$	$t = 2,3....$
X_1^2	X_1^1	X_t^1	$X_{t-1}^1 \quad F_{t-1}^1$	$t = 2,3....$
Y_t	$X_t^1 \quad X_t^2$	X_t^2	$X_{t-1}^2 \quad F_{t-1}^1 \quad X_t^1$	$t = 2,3....$

$$P\left(Y_t = y_t \mid X_t^1 = i, X_t^2 = j\right) = N\left(\mu_{i,j}, \Sigma_{i,j}\right)$$

Case Study: Machining Center (Drilling)

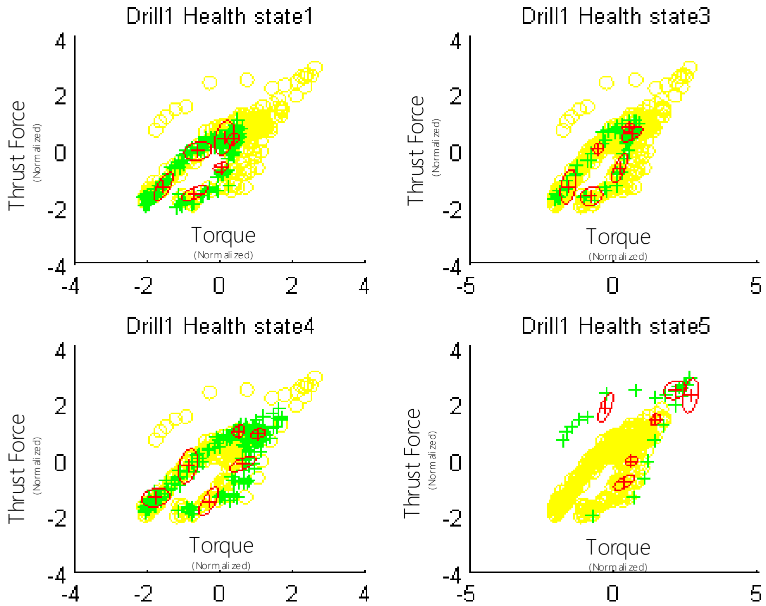
- **Goal**: Monitor the health-state of drill bits
- **Setup**: HAAS VF-1 CNC Machining Center
- **Sensor**: Kistler 9257B Piezo-Dynamometer
  - Thrust Force and Torque
- **Machining Conditions**:
  - Stainless Steel Plates: ¼" thickness
  - HSS drill-bits with two flutes
  - No coolant
  - Feed Rate: 4.5 ipm
  - Spindle Speed: 800 RPM
- **Thrust & Torque Data**:
  - 250 Hz
  - 380-460 data points per hole
  - Standardized to 24 RMS values



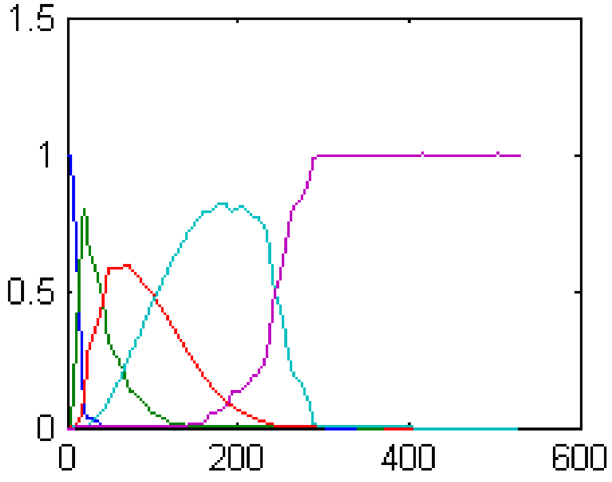


Hierarchical HMMs: Five Health States

HMM States



Likelihood



Health-state Estimation Results: All drill bits

Holes		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Drill bits	1	1	1	1	1	1	1	1	1	1	1	2	2	2	3	3	3	3	3	3	3	4	5		
	2	1	1	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	5						
	3	1	1	1	1	1	1	1	1	2	2	2	2	3	3	5									
	4	1	1	1	1	2	3	5																	
	5	1	1	1	1	1	1	1	1	2	2	2	3	5											
	6	1	1	[illegible]	[illegible]	1	2	3	5																
	7	1	1	[illegible]	[illegible]	1	1	1	2	[illegible]	[illegible]	2	3	3	5										
	8	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	4	5
	9	1	1	1	1	1	1	1	1	1	1	2	2	2	2	3	5								
	10	1	1	1	1	1	1	1	2	3	5														
	11	1	1	1	1	1	1	2	3	5															
	12	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	4	5							

Good Used Worn Fail

RUL Estimation: Monte-Carlo Simulation

- "Current" health-state information from diagnostics module
- Prognostics (Remaining-Useful-Life Estimation): Utilizing Monte-Carlo Simulations with Established Models
  - RUL: Number of transitions from the "current" state to the predicted "failure" state
    - Current Setting: Number of holes to be successfully drilled by drill-bit
  - RUL distribution from several Monte-Carlo runs
    - RUL mean and confidence intervals

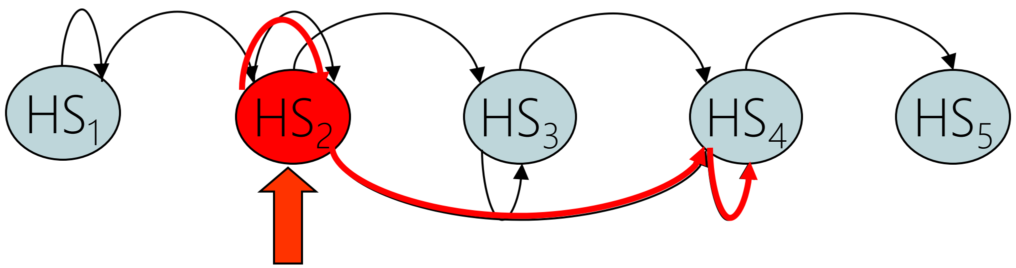
ILLUSTRATIVE EXAMPLE:

Health state transition probabilities:

$p_{2,2} = 0.4$

$p_{2,3} = 0.5$

$p_{2,4} = 0.1$



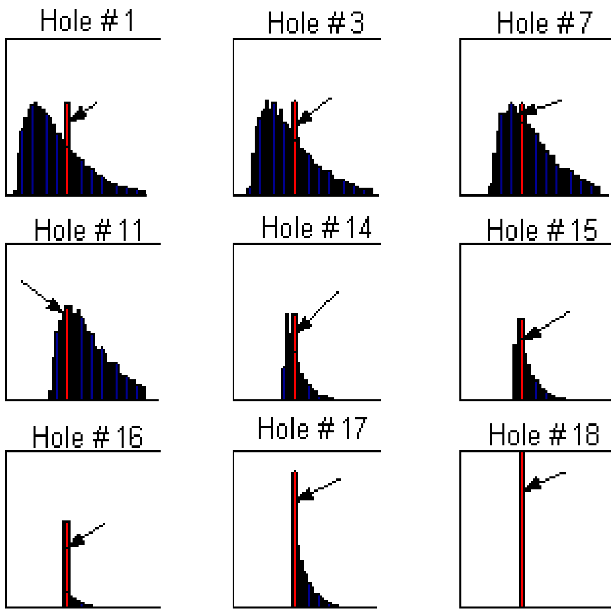
RUL Samples:

5, 5, 3, 4, 5, 3, ...

RUL Estimation: Results

Drill Bit #18

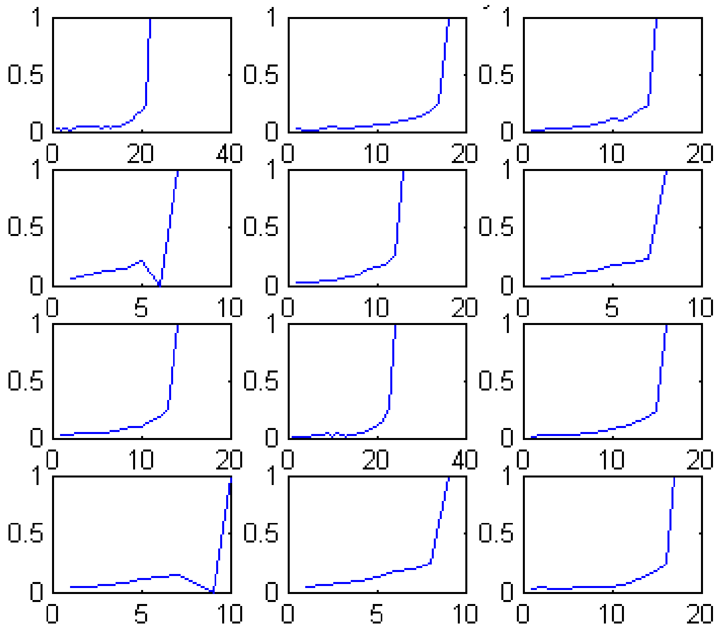
RUL Probability Distribution



Drill-bit Life (Holes)

All Drill Bits

Estimation Accuracy



Frequency with which actual RUL is within estimated RUL confidence limits

Monitoring & Diagnostics of Industrial Equipment at Scale

Clustering and Cluster Tracking Agents

IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 6, NO. 4, NOVEMBER 2010 767

An Industrial Strength Novelty Detection Framework for Autonomous Equipment Monitoring and Diagnostics

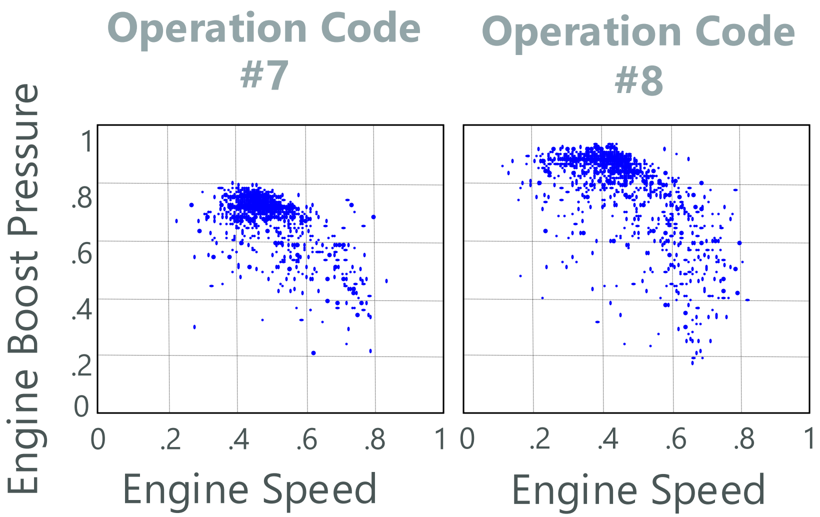
Dimitar P. Filev, Ratna Babu Chinnam, Finn Tseng, and Pundarikaksha Baruah

Novelty Detection Difficulty: "Operating Modes"

- How to recognize (compare numerically) differences between different asset/machine "operating modes" (regimes)?
  - Function of system, operating condition/load, and the external environment
- How to distinguish "normal" machine operation from a "fault" condition?

Hydraulic Excavator of Shin-Caterpillar Mitsubishi

	Code	Label	Operation Description
FAULTY CONDITIONS	1	F1	Fuel Spray Nozzle deactivated
	2	F2	Turbo-Charger Deterioration
	3	F3	Valve Clearance Changed
	4	F4	Air Filter Obstruction (High)
	5	F5	Air Filter Obstruction (Low)
NORMAL	6	NH	Operation (High Load)
	7	NM	Operation (Medium Load)
	8	NL	Operation (Light Load)

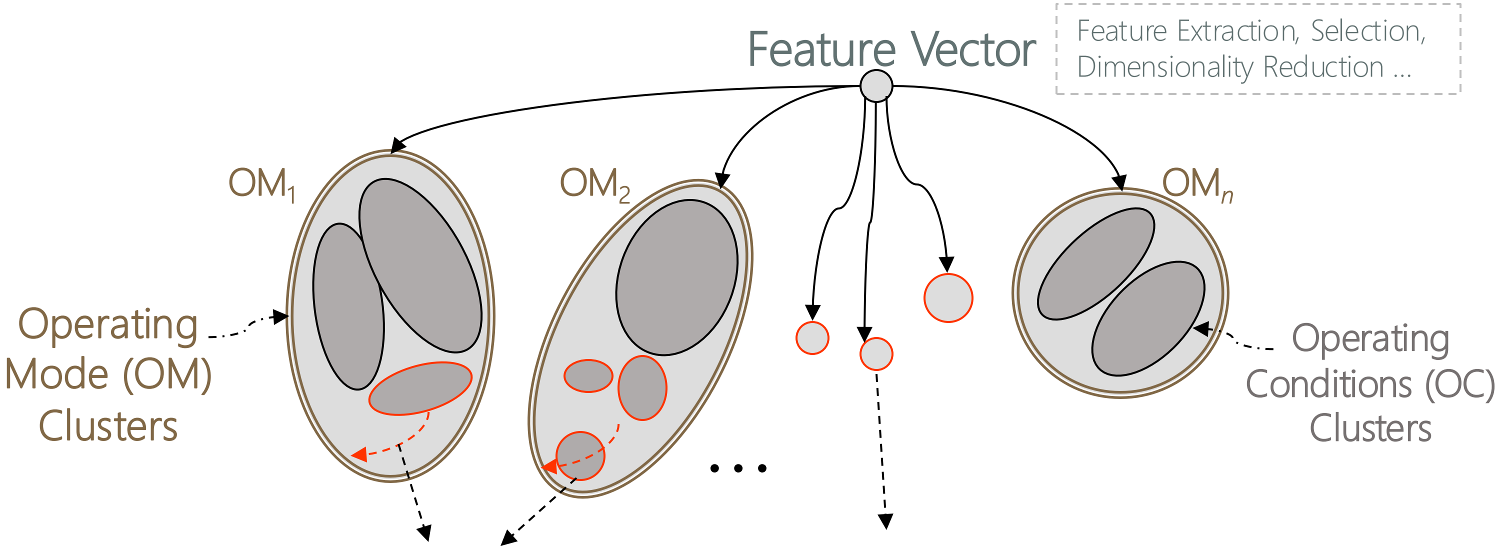


(Source: Dimitar Filev 2007)

Detecting Modes: Clustering & Tracking Agents

- Clustering methods can be effective at detecting and tracking operating modes
- Can learn from data (structure/parameters)
- Can adapt to a changing environment
- Summarization and decision-making ability

DETECTABLE OPERATING MODES:
Dictated by the combination of Target System type, Sensors, Data Acquisition Scheme, and Extracted Features

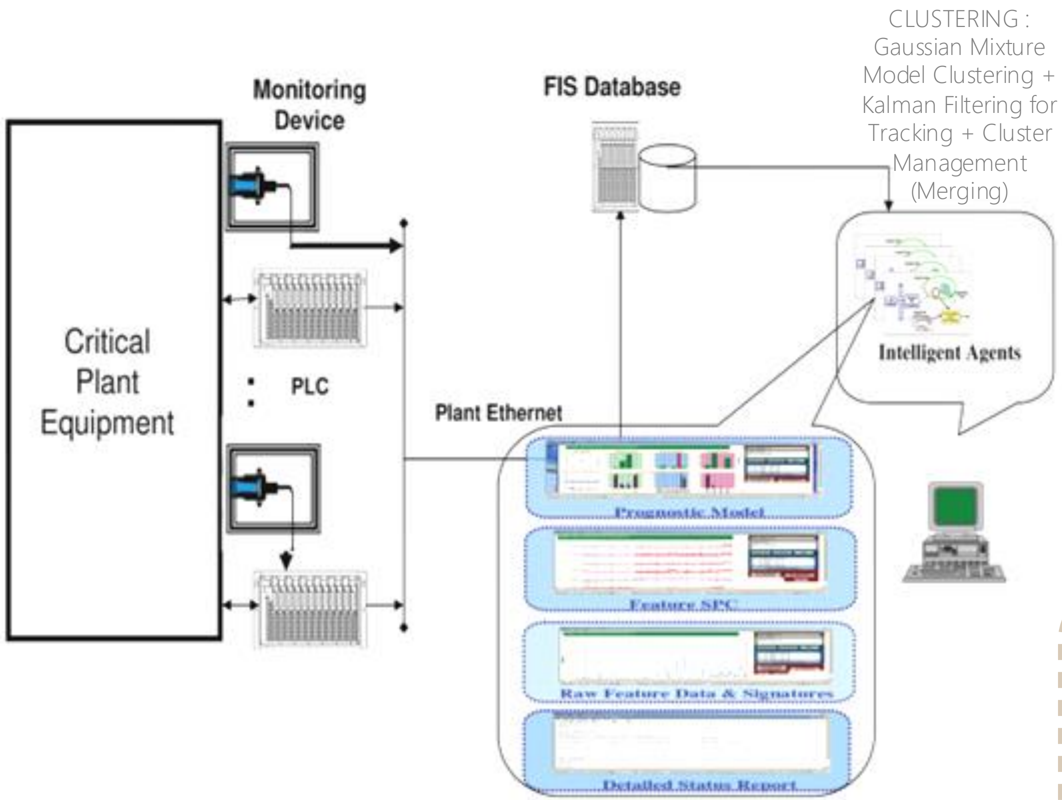


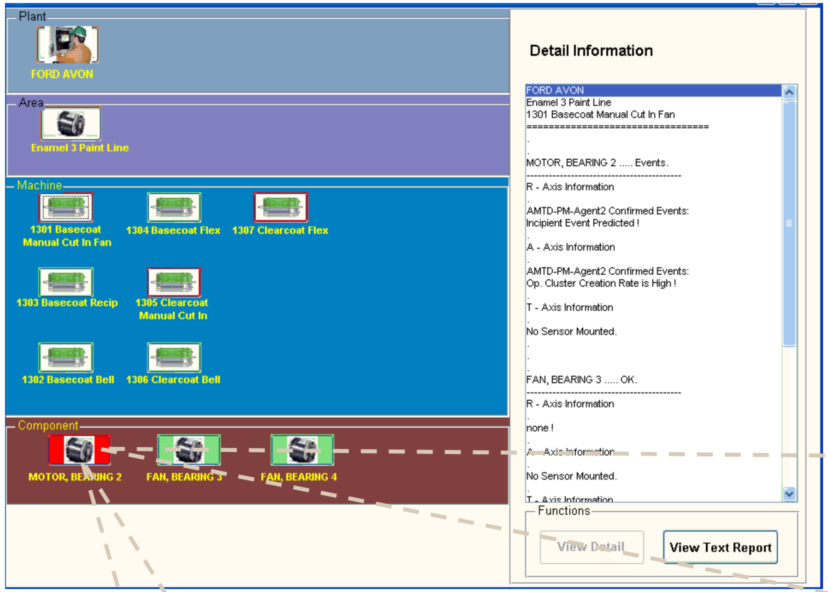
Fault Type I (Incipient): *Cluster centers approaching boundary of OM cluster*

Fault Type II (Drastic): *Rapid generation of new OM clusters.*

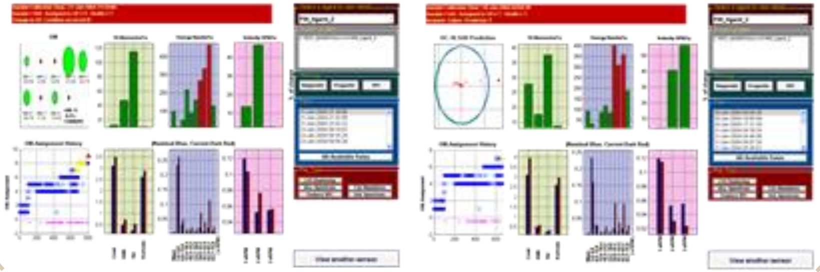
Scalable Platform for Diagnostics & Prognostics

Cost-Effective Client-Server Solution for Manufacturing Plants



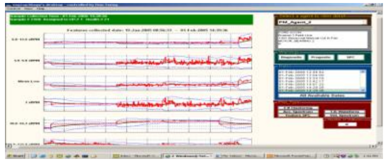


PREDICTED FAULT WARNING COMBINED WITH A SUMMARIZED STATUS OF FEATURES

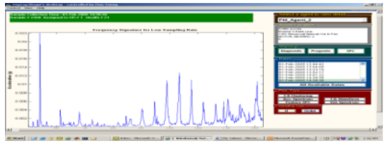


SUMMARIZED MACHINE HEALTH REPORT

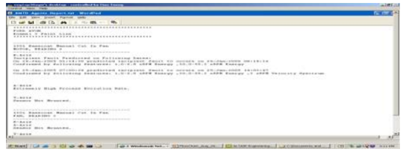
OM & OC Conditions



Feature SPC



Raw Data & Freq. Signatures



Detailed Status Text Report

(Source: Dimitar Filev 2007)

Connected Vehicle Prognostics

Clustering and Cluster Tracking Agents

A mutual information based online evolving clustering approach and its applications

Fling Tseng, Dimitar Filev & Ratna Babu Chinnam

Original Paper | 15 July 2017 | Pages: 179 – 191

Brake Pad: Wear Prognostics

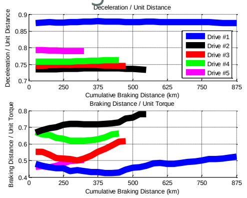
Setting: Vehicle CAN Bus Data; 5 Vehicle Fleet with Regular Brake Pad Inspection

Connected Vehicle

Sensor Signals →

- CAN Bus Signals
- Cumulative usage patterns
- Generic vehicle information (e.g., odometer mileage)

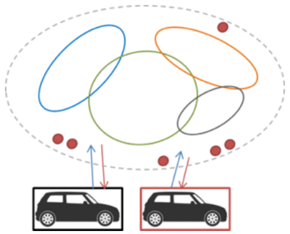
Braking Preference



Braking Performance

Feature Extraction →

- Braking preference
- Braking efficiency
- Braking conditions



Evolving Clustering →

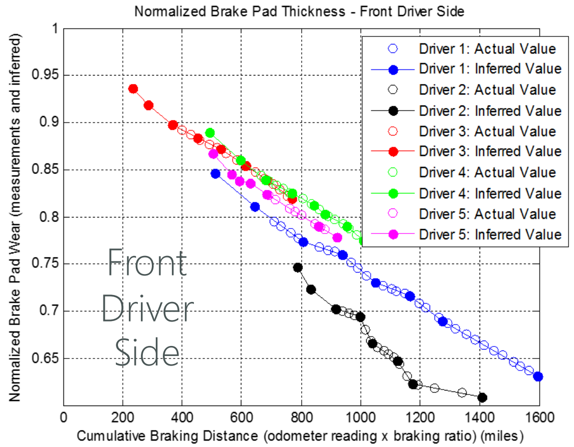
- Evolving Clustering (MIRGKL)

Occasional Brake Wear Measurement

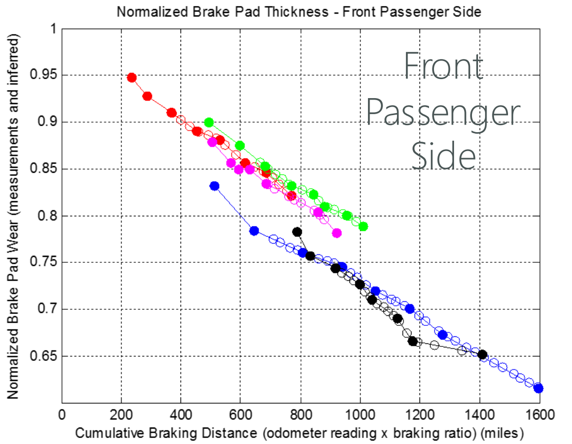
True Condition

Prognostics

- Learning of Multiple Degradation Models



Brake Wear vs. Cumulative Braking Distance



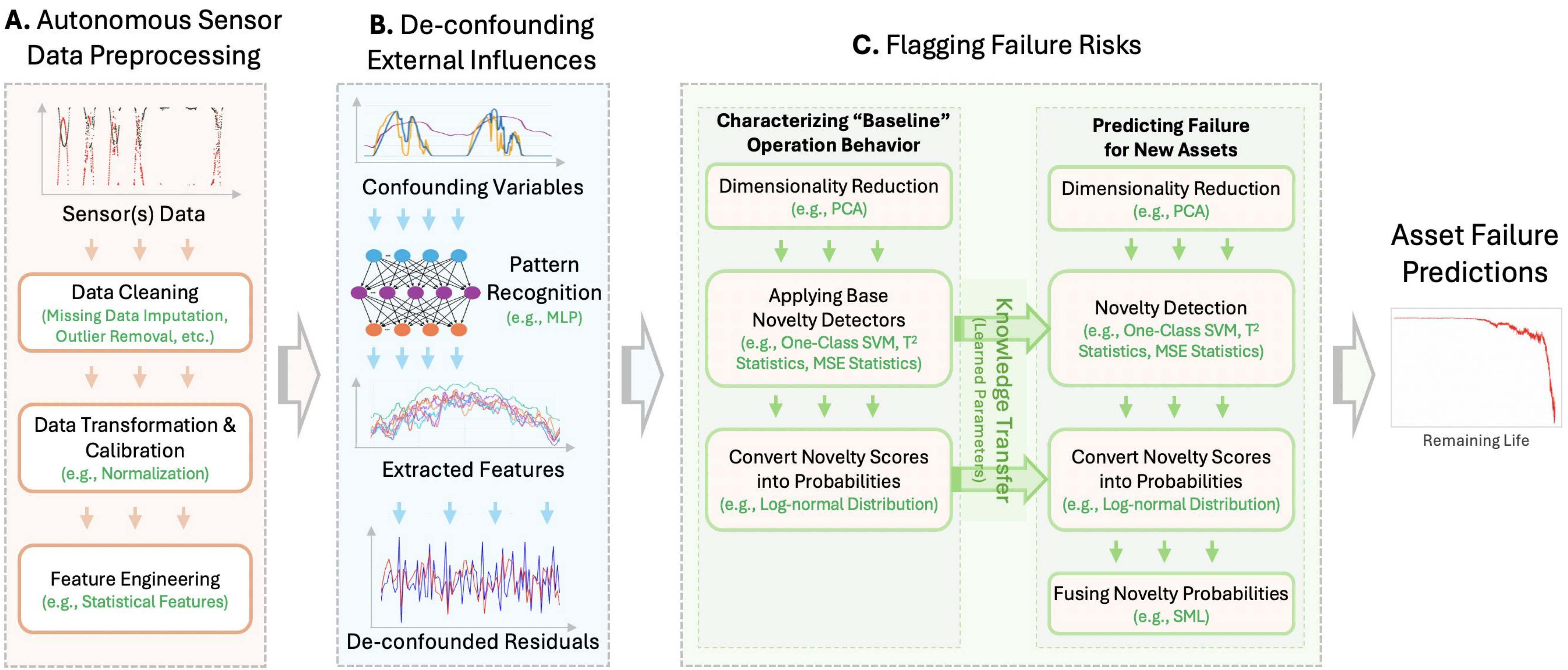
LIMITATION: Required feature engineering and consultation with SMEs!

Novelty Detection for Energy Asset Monitoring

Framework for Advancing Autonomous Monitoring & CBM

Preemptive Failure Prediction Framework

Monitoring and Novelty Detection



Proposed Ensemble Algorithm

Monitoring and Novelty Detection in Energy Assets

- Convert novelty scores to probabilities
  - Employed Lognormal distribution for effectiveness
  - Employed a probability threshold to produce a binary output for each detector
- Fuse binary outputs
  - Spectral Meta Learner (SML) used for fusion
  - Proposed Robust SML to enhance robustness

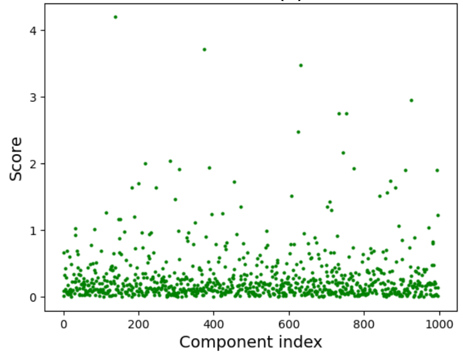
$$\hat{Y}_j^{SML} = sign(\sum_i f_i(x_j)\hat{V}_i)$$

Health state of asset j

Binary label of base detector i for component j

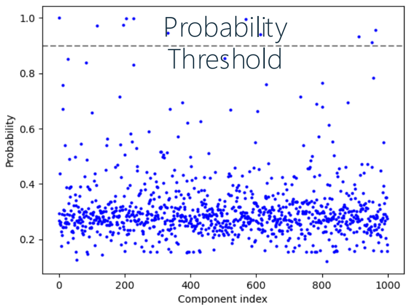
Dominant eigenvector of population covariance matrix of base detectors

Base Detector(s) Scores



Convert novelty scores into probability

Novelty Probabilities



Applying Probability Threshold

Asset	Detector 1	Detector 2	...	Detector n
1	0	0	...	0
2	0	1	...	1
...	...	...	...	...
m	1	0	...	0

Fusing base detectors output

Asset Health State

Asset	Health State
1	Normal
2	Anomalous
...	...
m	Normal

Case Study: 12-Volt Batteries (Fleet Trucks)

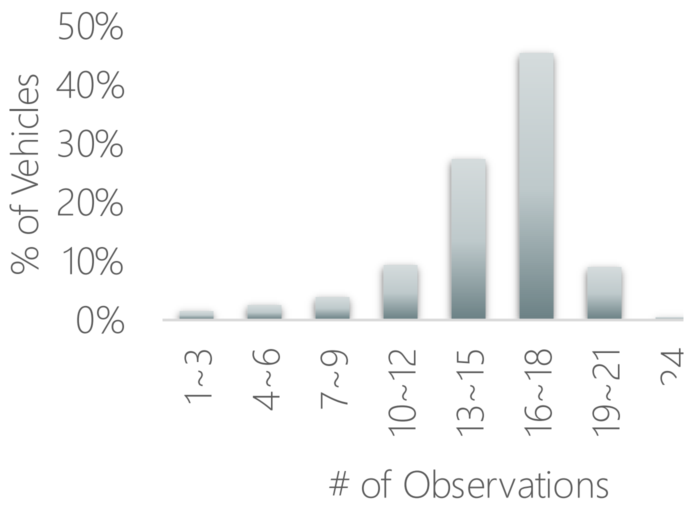
Novelty Detection Framework for Monitoring Connected Vehicle Systems with Imperfect Data

Manuscript Under Review

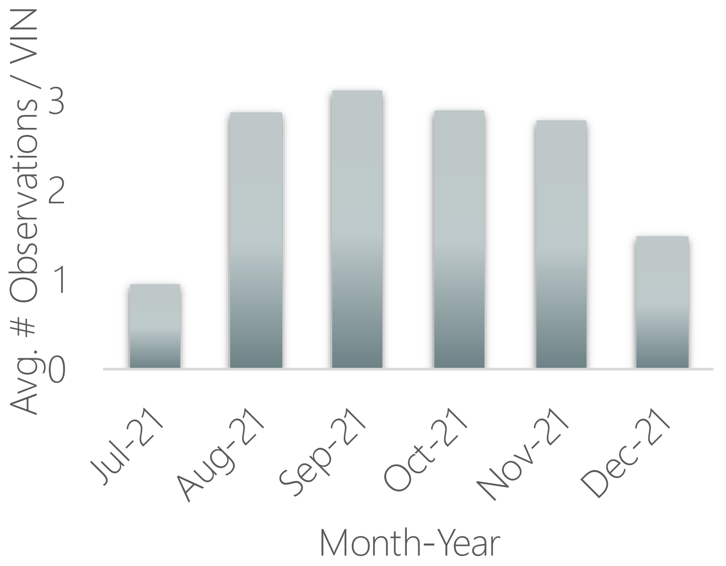
Case Setting: 20k Vehicles, 50 US States (Large OEM)

Sparse & Erratic Sampling

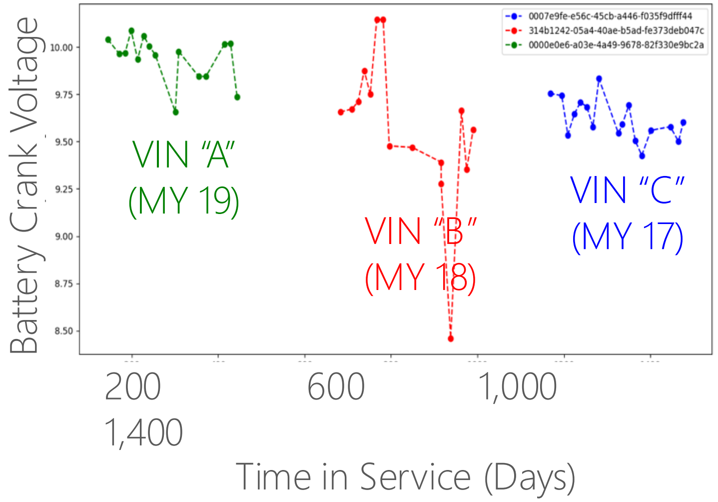
- Samples per VIN



- Avg. # Samples by Month



- Sparse & Uneven History



- Example Vehicle: Three sparse observations

Observation Time	Crank Current	Crank Voltage	Current Flt	OCV	Capacity	Resistance	SOC
2010-08-11 13:50:00	-1058	9.445313	28.875	12.8125	79	3.3125	70
2021-10-04 23:12:56	-1102	8.402344	35.59375	12.8125	73	3.3125	39
2021-12-14 19:52:45	-1004						

Novelty Detection Results

A "Vehicle Model & Engine" Combination

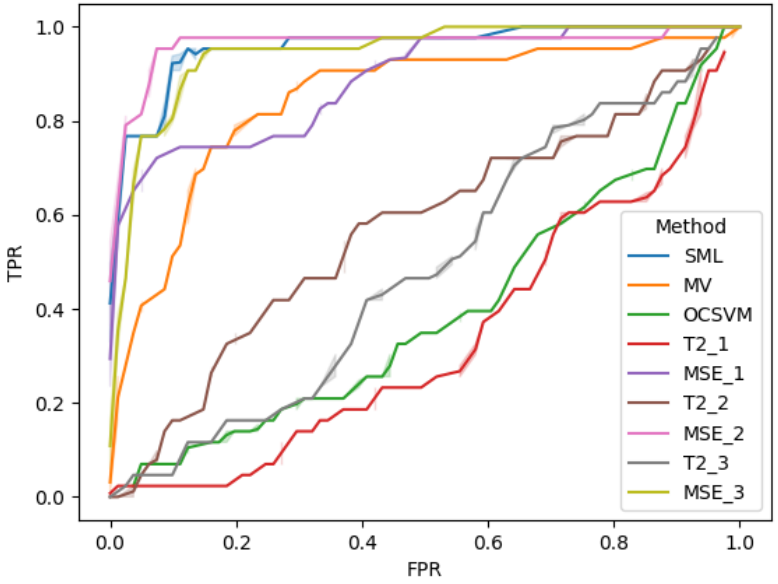
- **Framework demonstrated strong performance** in detecting anomalous batteries.
- **Effectively handled noisy and incomplete sensor data**, with expert validation confirming the relevance of many flagged anomalies.
- **Robust SML method outperformed traditional approaches.**

Avg. Lead Time for Detection:
8.6 Days ahead of Battery Reset

	Identified to be "Normal"	Identified to be "Anomalous"
22 Vehicles with Battery Reset	3	19
Vehicles without Battery Reset	6,584	63

Batteries might have been reset for other reasons

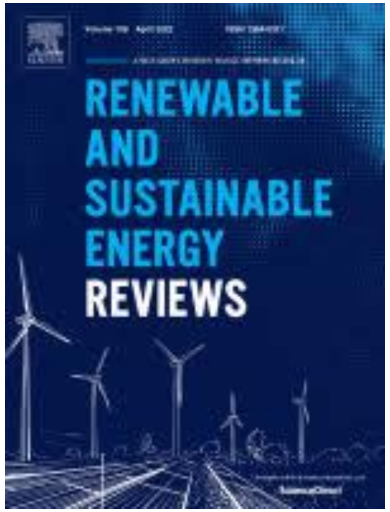
Good sensor data evidence that some of these batteries might indeed be faulty



ROC curves for individual detectors, majority voting, and the SML algorithm

Case Study: Photovoltaic Inverters

A Modular Framework for Sensor-Driven Failure Prediction in Energy Systems: An Industrial Case Study of PV Inverters

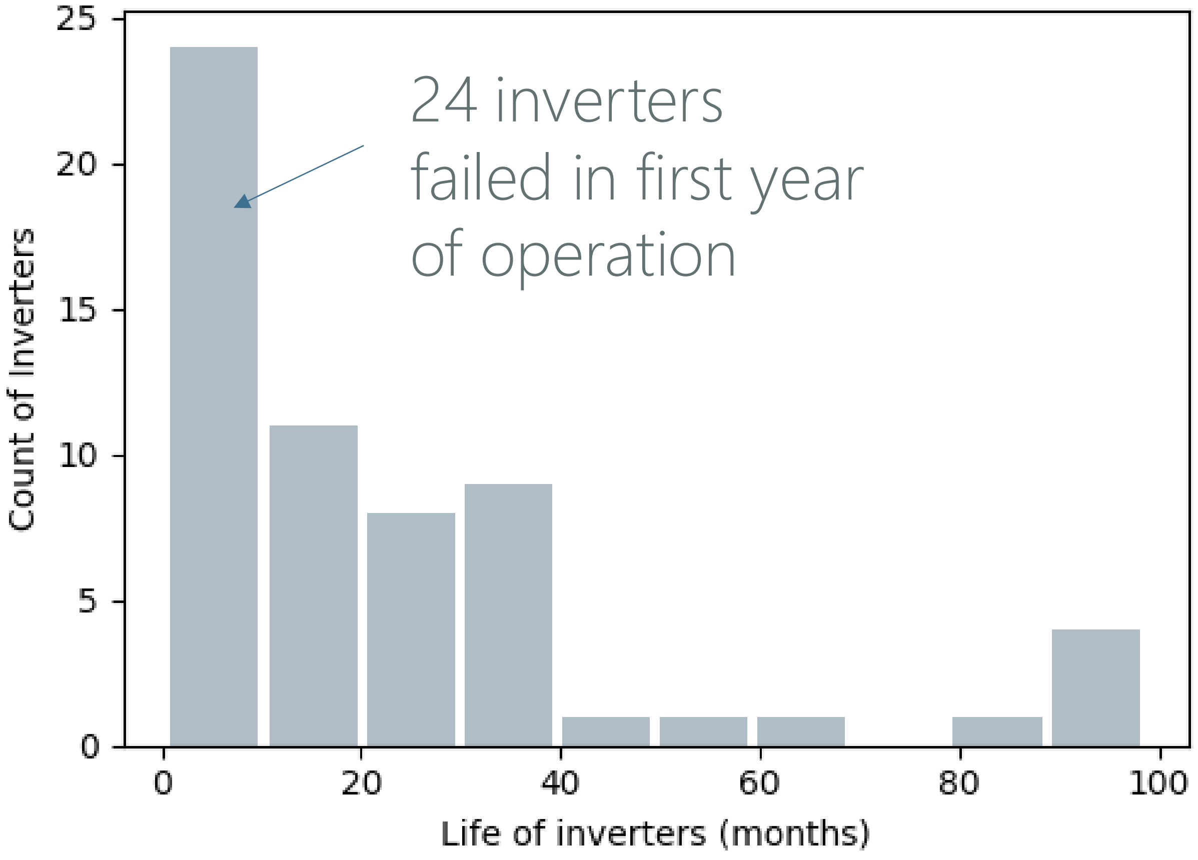


Manuscript Under Review

Case Setting: 59 Inverters, Different Models/Regions

Sparse & Erratic Sampling

- **Period**: 2014 to 2022
- **Location**: Different sites/states
- **Sensors**:
  - AC Power
  - AC Voltage
  - DC Voltage
  - AC Current
  - AC Frequency
- **Data Resolution**: 5 minutes

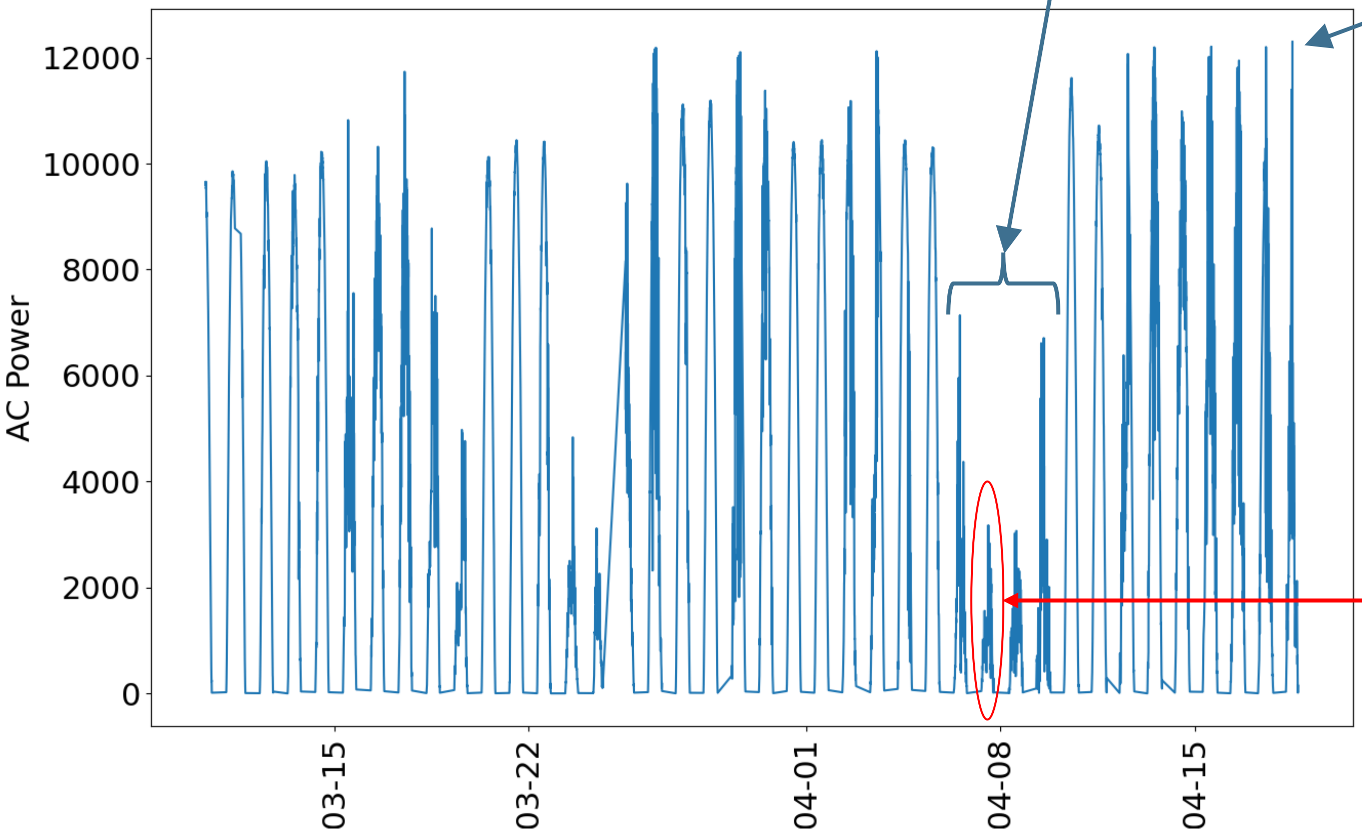


Data Description

Single Inverter History

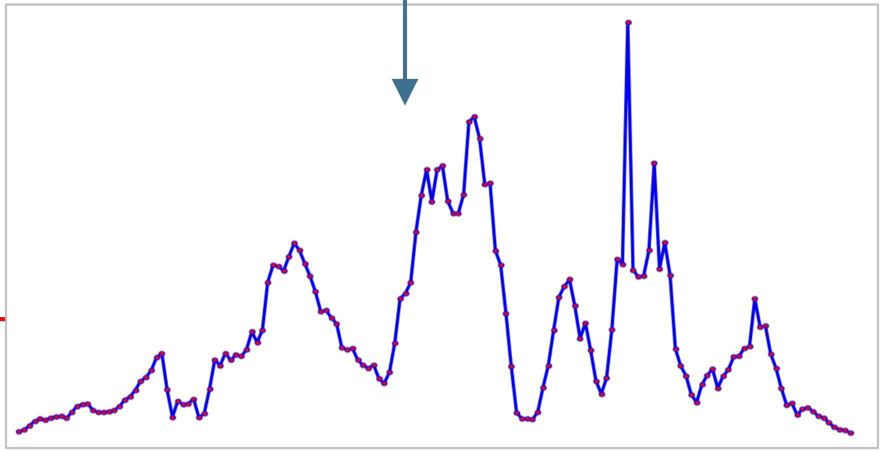
Is this downturn indicating an anomaly?

No obvious trend of abnormal behavior prior to failure



AC Power measurements for an inverter **until failure**

Significant fluctuation. **Why?**



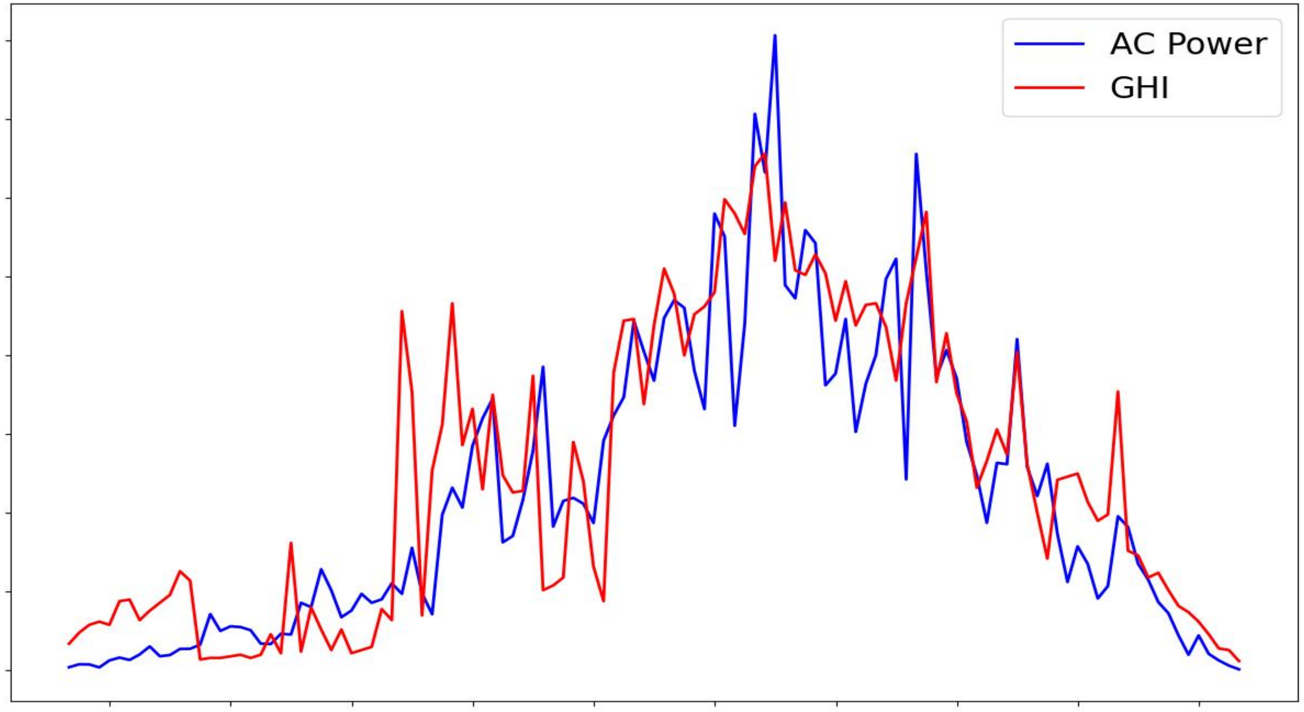
AC Power measurements for **one day**

5-minutes resolution

Data Description

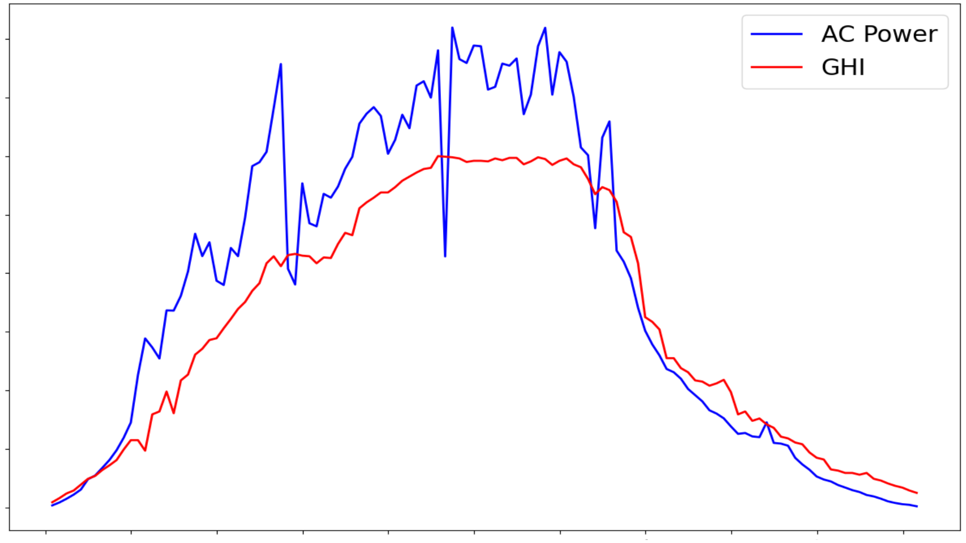
Root Cause for Fluctuation: Environment (solar radiation)

Correlation between solar radiation and AC Power



AC Power measurements for **one day**
5-minutes resolution

GHI: Global Horizontal Irradiance
is an index for solar radiation

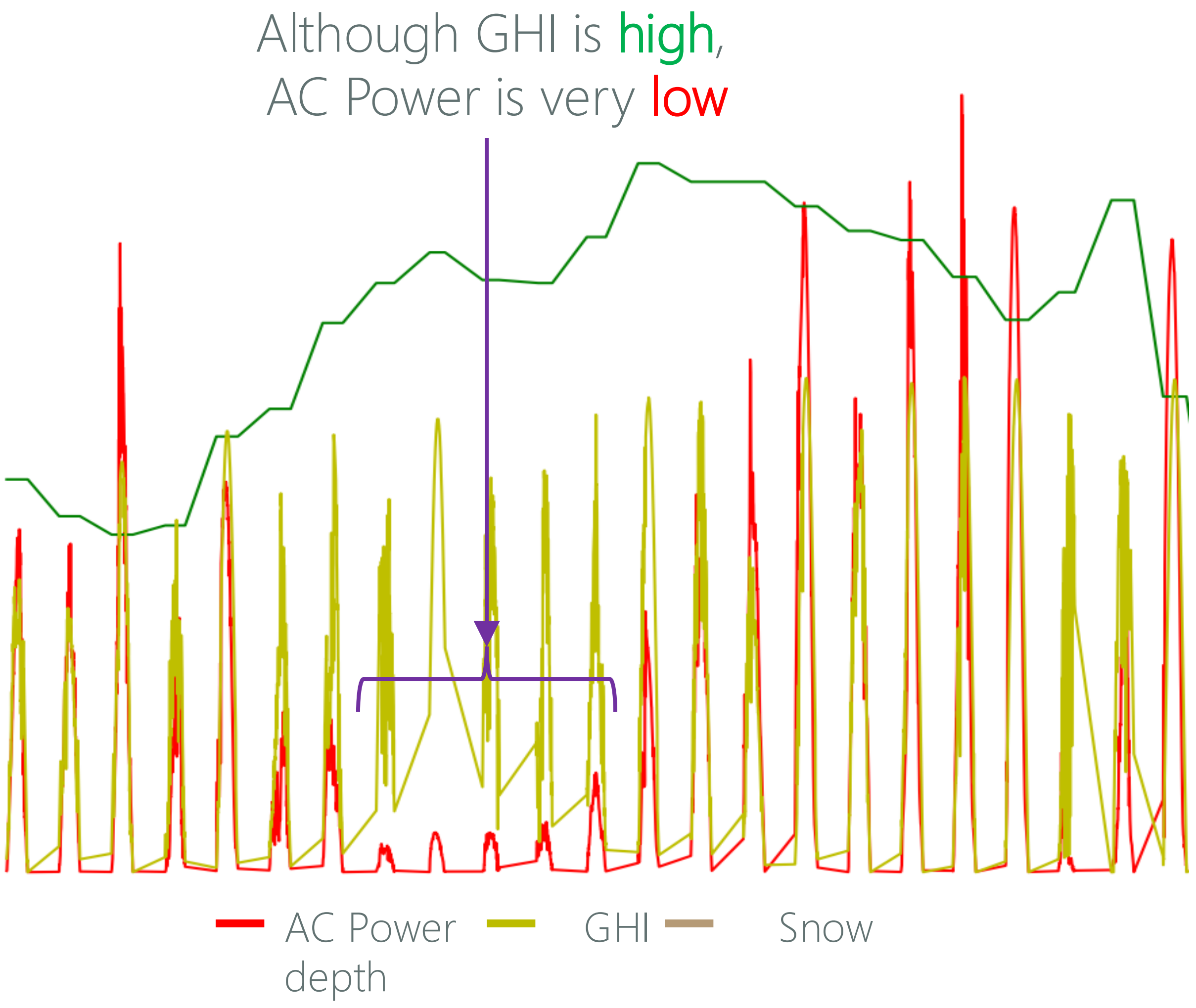


Data for another day!

Data Description

Another Environmental Cause for Fluctuation: Snow Cover

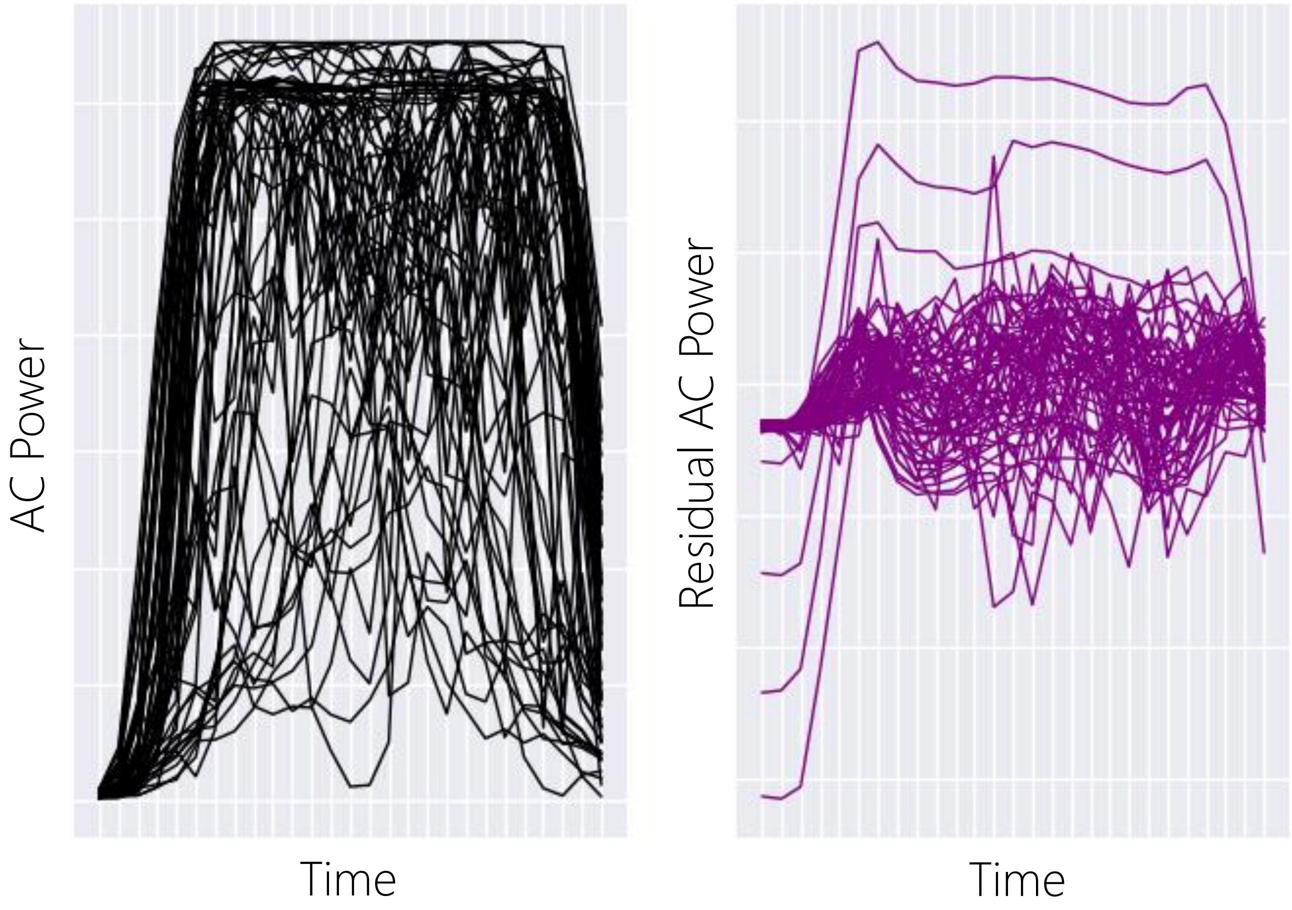
- **Snow covers solar panel** surface and blocks solar radiation from reaching the panels
- **Manual snow removal** or melting resumes normal situation
- **No data on removal of snow**
- Analyzing **abrupt shifts in AC Power** on snowy days, we inferred snow cover.



Deconfounding Influences

Multi-layer Perceptron Model

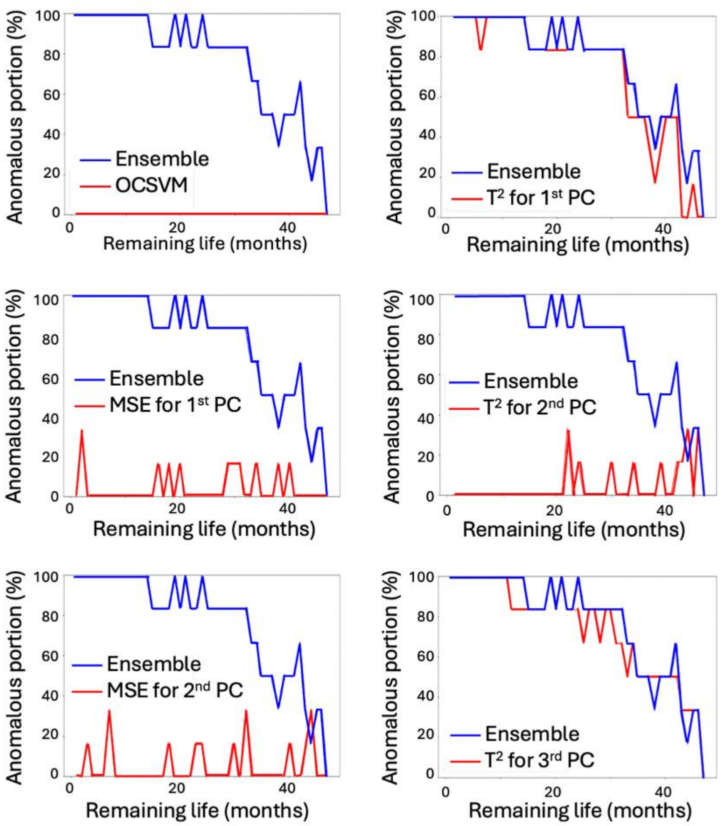
- Employed multi-layer perceptron (MLP) to regress relationship between environmental variables and sensor measurements
- Residuals of regression model represent deconfounded sensor measurements
- 74% of AC Power variance explained by model



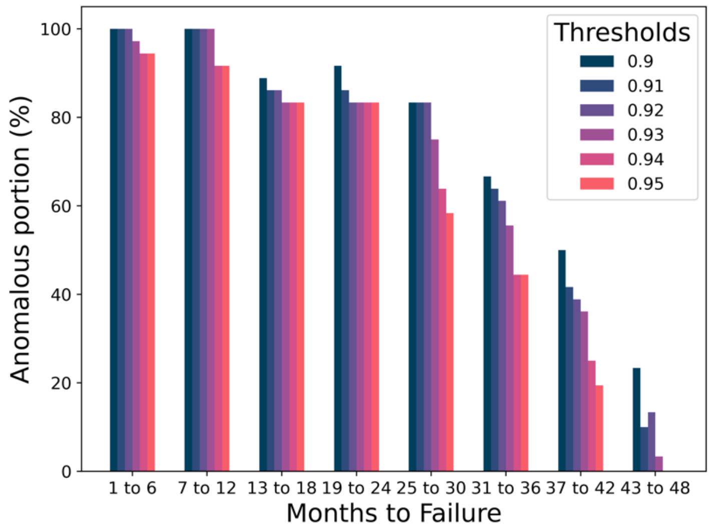
Novelty Detection Results

Photovoltaic Inverters

- Framework successfully identified PV inverters approaching failure.
    - Flagged anomalies in the months leading up to failure with decent accuracy
- Results demonstrate robustness in handling noisy and confounded sensor data.
- Robust SML method outperformed individual detectors, reducing false positives and improving the accuracy of failure predictions.



% of PV inverters detected as anomalous over their remaining life



% of PV inverters detected as anomalous over their remaining life with varying novelty score thresholds

Novelty Detection

Deep Learning Models for Warranty Issue Detection

Assembly -> Customer -> Claim

Deep Learning for Warranty Quality Issue Detection

Traditional Monitoring

- Relies on tolerance checks at assembly stations and end-of-line tests.
- Testing does not fully represent real-world driving conditions.
- Fails to capture interactions across different assembly stations effectively.

Proposed Approach

- Developed a neural network model to detect patterns linking IIoT data across stations and warranty claims.
- Utilized transfer learning to adapt the model efficiently to new vehicle models with minimal additional data.

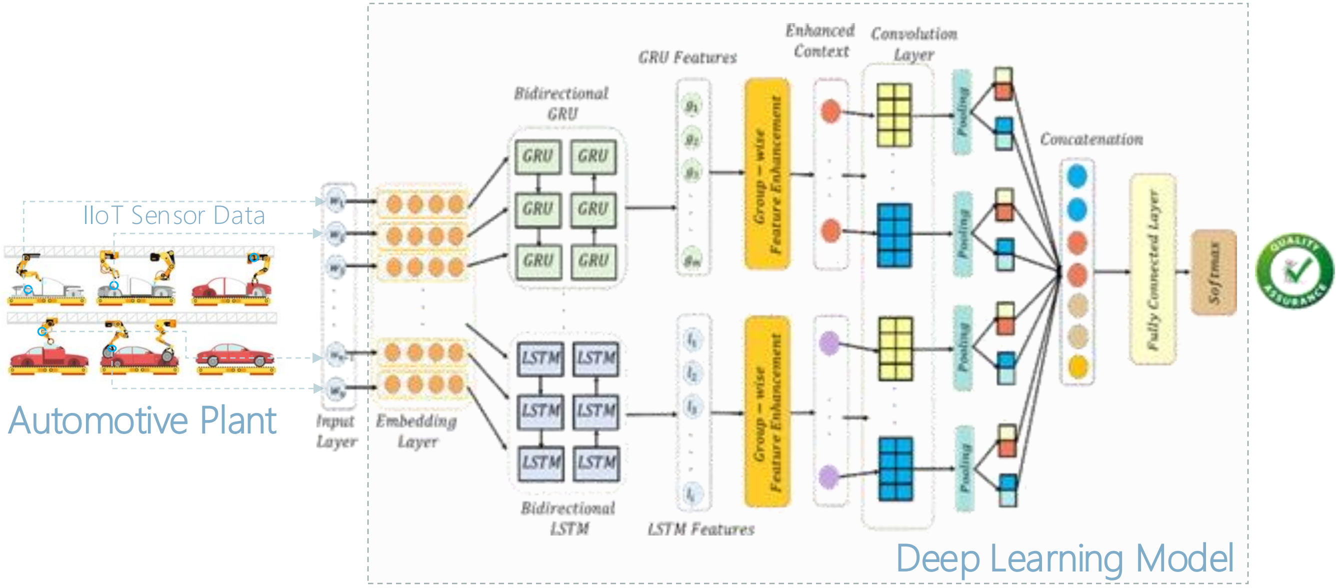


Photo Credits: Aytuğ Onan and Others

Case Study: Large OEM Assembly Facility

- **Vehicle**: Four-Door SUV
- **Dataset**: 64,774 Vehicles (Jun- Nov 2022)
- **Warranty Codes & Claims**: Steering Wheel Alignment
  - Type-1: Pulls (193); Type-2: Off-Center; Type-3: Wander
- **Industrial IoT Setup**:
  - 62 Stations with 392 Sensors, > 25 million measurements
- **Promising Results**:
  - Identified 60% of vehicles with Pulls claims; few false positives.
  - Model demonstrated effective transfer to two-door Wrangler models despite limited data.
- **Transfer Learning**: Of 3,087 two-door SUVs, 8 of 13 with claims were identified; 2 potential false positives.
  - Unsupervised domain adaptation using CORAL

MRI Anomaly Detection

Variational Auto-Encoders

nature scientific reports

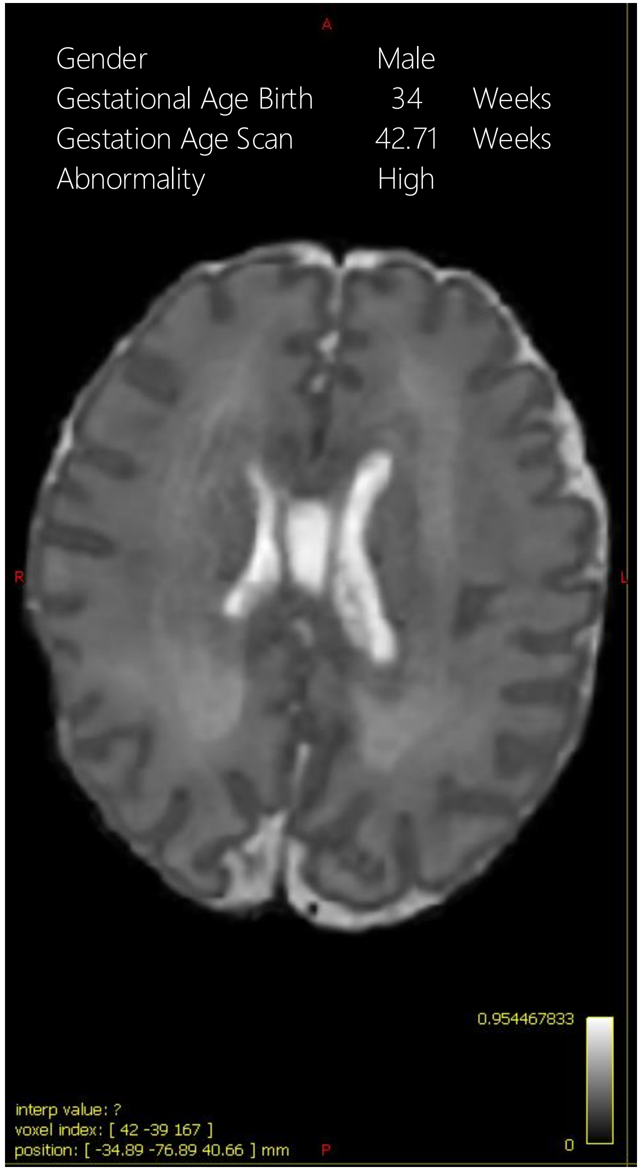
Article | Open access | Published: 17 July 2023

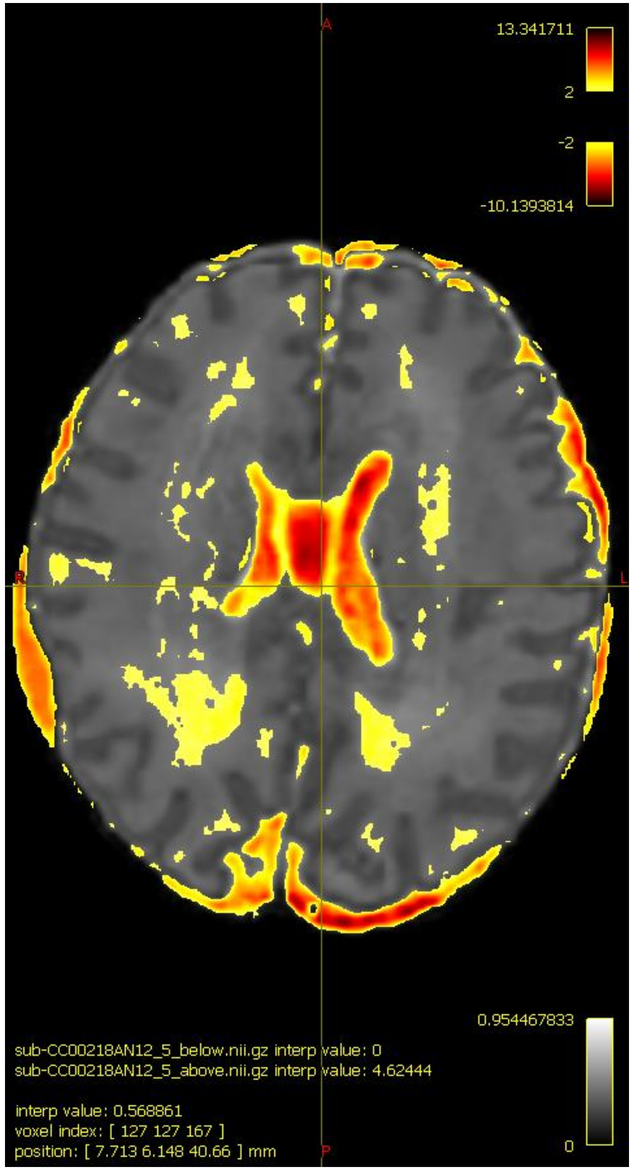
Unsupervised abnormality detection in neonatal MRI brain scans using deep learning

Jad Dino Raad, Ratna Babu Chinnam, Suzan Arslanturk, Sidhartha Tan, Jeong-Won Jeong & Swati Mody

Variational Auto-Encoders for MRI Anomaly Detection

- **Setting**: Neonatal MRI Images
- **Data Source**: Developing Human Connectome Project (dHCP)
  - King's College London, Imperial College London and U. of Oxford
- **Collaboration**: Wayne State University & Children's Hospital of Michigan
  - WSU: Jad Raad, & Drs. Chinnam, Arslanturk
  - CHM: Drs. Tan, Mody, Jeong
- **Focus**: Automated Anomaly Detection
- **Approach**: Deep Learning Variational Auto-Encoders
- **Preliminary Results are Very Promising!**





Conclusion

Conclusion

- **Current Reality**: Despite decades of R&D, monitoring and diagnostics still in its infancy.
- **Scaling Challenges**: Widespread adoption requires further advancements in fully autonomous methods.
- **Turn-key Solutions**: Future systems must self-calibrate, self-learn, and adapt autonomously to dynamic conditions.
- **Deep Learning Advantages**: Reduces dependence on manual feature engineering, enabling greater automation.
- **Key Areas for Growth**:
  - **Transfer Learning –** Expanding model adaptability across diverse asset types and applications.
  - **Federated Learning –** Enhancing model robustness while preserving data privacy.
  - **Hybrid Methods –** Integrating domain knowledge for more reliable solutions.
- **Accelerating Innovation**: Open-source contributions, expanded datasets, and collaborative GitHub repositories can drive rapid advancements.

Collaborators & Credits

PhD Students & Graduates:

- Pundarikaksha Baruah, PhD
- Fatih Camci, PhD
- Akhilesh Kumar, PhD
- Fling Tseng, PhD
- Mohammad Badfar, PhD Candidate
- Sujeet Shrestha, PhD Candidate

MS Students:

- Vinodh Rengasamy
- Pankaj Kumar
- Shubha Murthy
- Asif Yasin
- Alok Rathod

Faculty Collaborators:

- Murat Yildirim, PhD
- Alper Murat, PhD
- Suzan Arslanturk, PhD

Industry Collaborators:

- Dimitar Filev, PhD, Ford (Retired)
- Chongwen Zhou, PhD, Faurecia
- Stellantis
- Drs. Sid Tan, Swati Mody, CHM
- Alper Unler, PhD, Turkey
- Argonne National Lab

Manufacturing Engineering Building
Home of Industrial & Systems Engineering Department

THANK YOU

Ratna Babu Chinnam, PhD

ratna.chinnam@wayne.edu

Marvin Danto Engineering Development Center
College of Engineering