Automated Ultrasonic NDE Data Analysis

Hongbin Sun, Muthu Elen, Pradeep Ramuhalli, Ryan Meyer, Matt Prowant, Richard Jacob

NRC/Industry NDE Technical Information Exchange Meeting

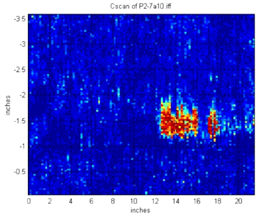
Jan 22-23, 2025

This work was sponsored by the U.S. NRC
Carol Nove, NRC COR

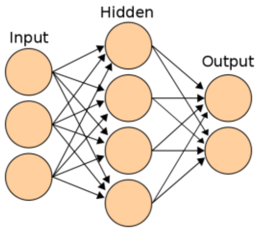
ORNL IS MANAGED BY UT-BATTELLE LLC AND PNNL IS OPERATED BY BATTELLE FOR THE U.S. DEPARTMENT OF ENERGY

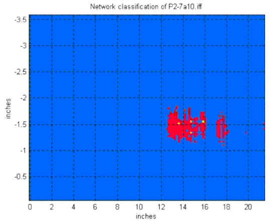
ORNL Pub. ID: 228827
PNNL-SA-207594

Objective: Assess current capabilities of machine learning (ML) and automated data analysis for improving NDE reliability

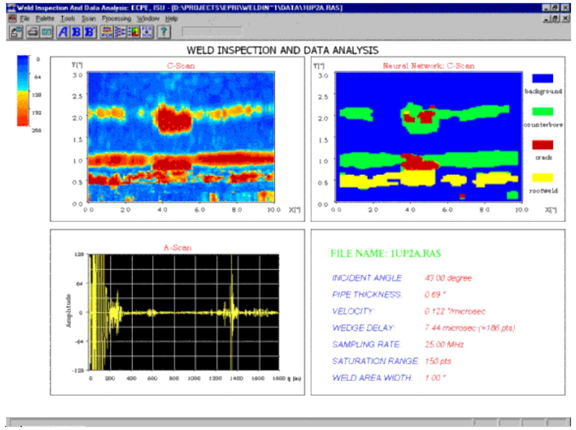


Single-element UT NDE Data





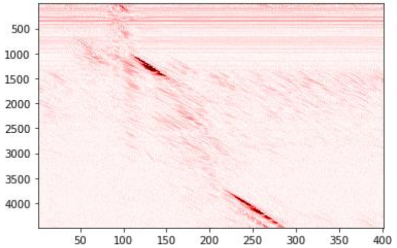
NN Classification Result

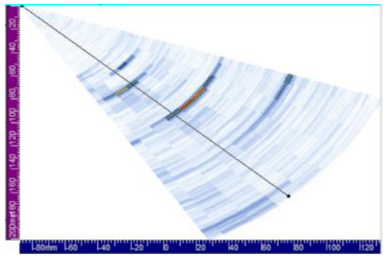


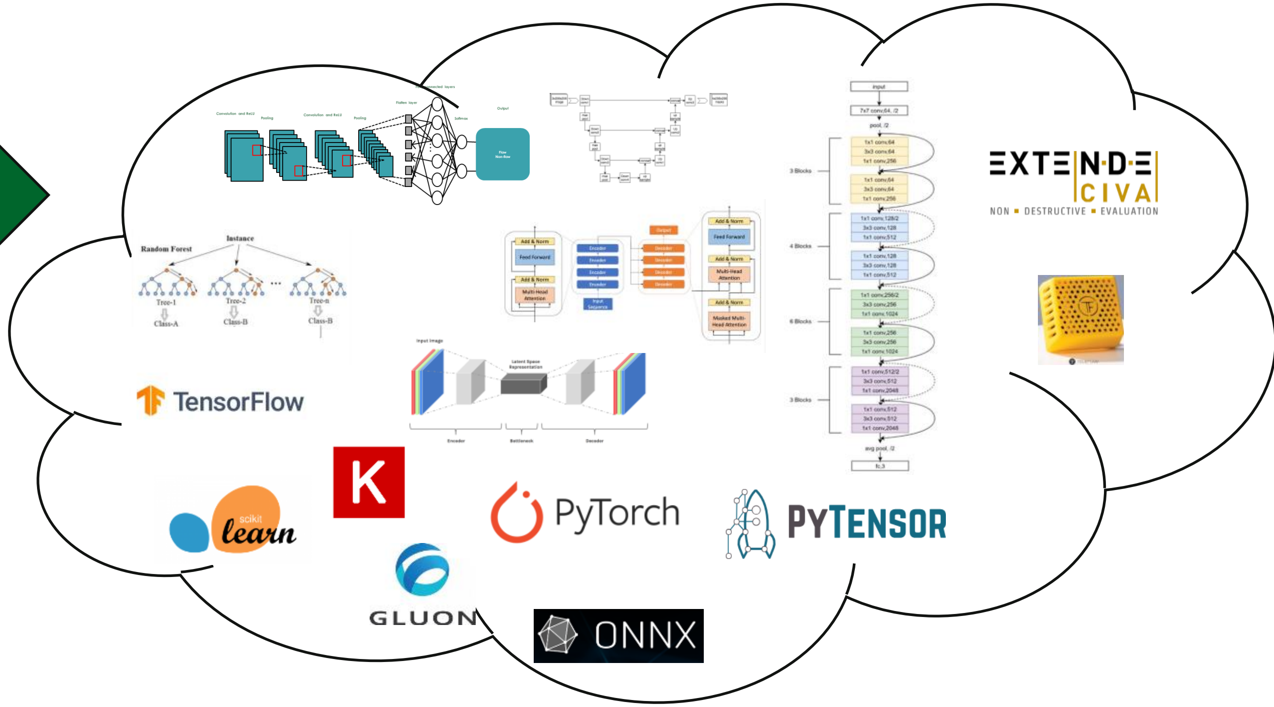
Examples of Machine Learning Applied to IGSCC (Examples from open specimens)*

*Spanner et al, 2nd Int'l. Conf. NDE in Relation to Structural Integrity for Nuclear and Pressurized Components, New Orleans May 2000

ML for UTNDE Data Analysis – circa 2000

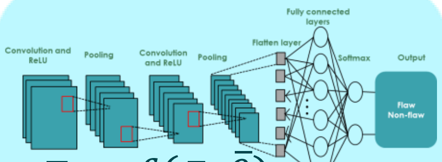






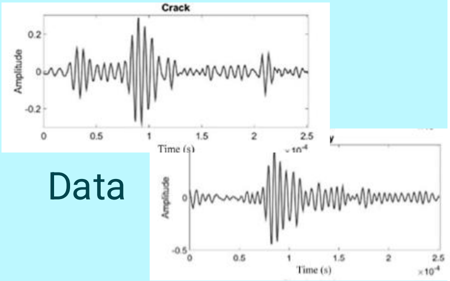
UTNDE and ML Landscape – circa 2025

Objective: Provide technical basis to support regulatory decisions and Code actions on automated data analysis for NDE



$\bar{y} = f(\bar{x}, \bar{\theta})$

Model and Model Parameters

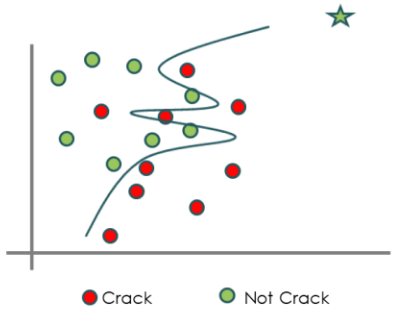


Data

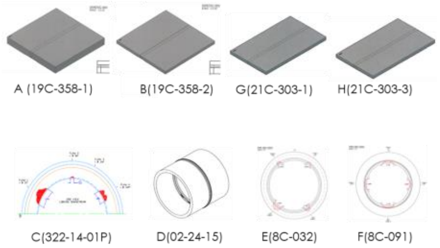
https://www.zetec.com/blog/destructive-and-nondestructive-testing-of-welds-how-ndt-ensures-quality/

Automated Analysis/AI/ML Purpose: Application

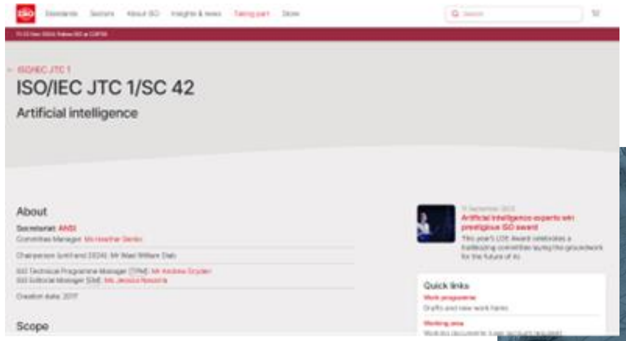
Drivers Influencing ML Performance

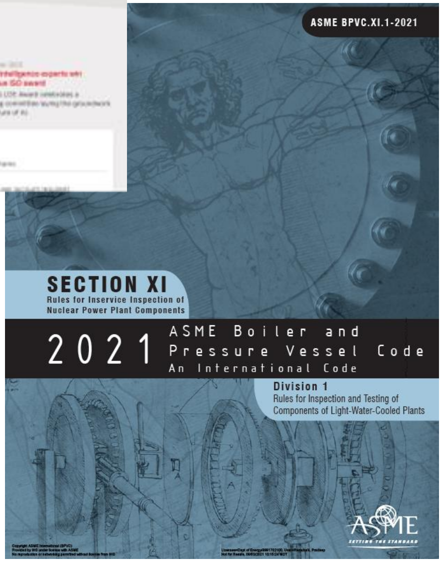


Validation and Qualification Requirements?



Data Requirements?

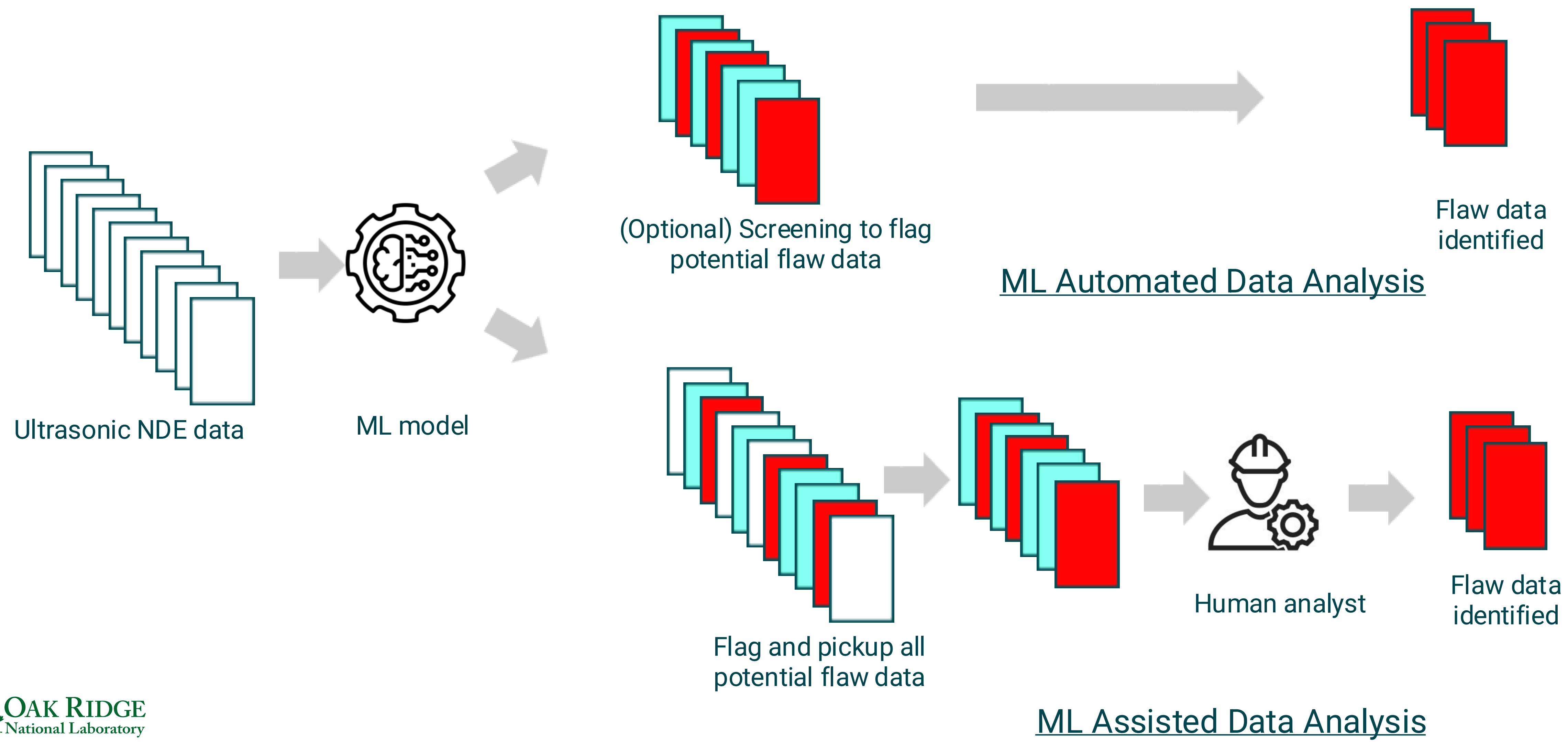




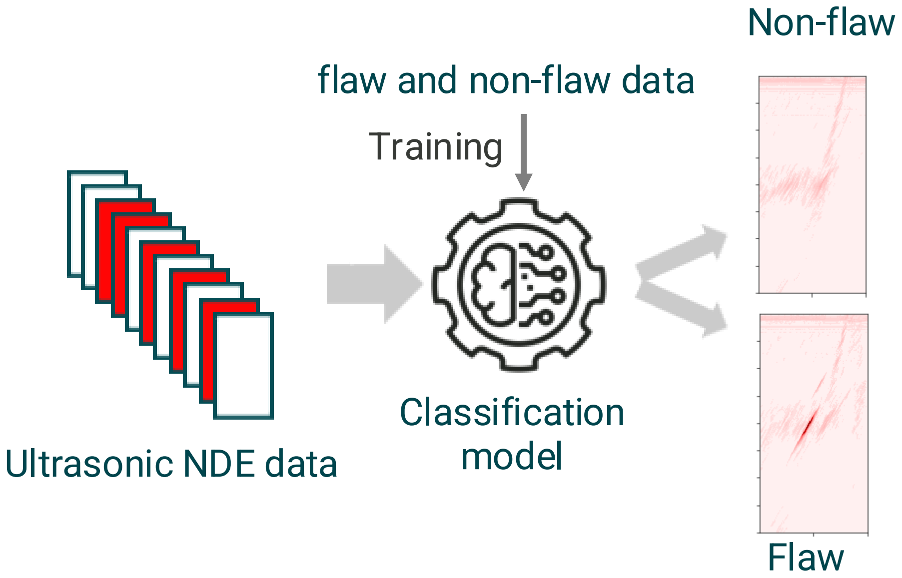
Codes and Standards?

OAK RIDGE National Laboratory

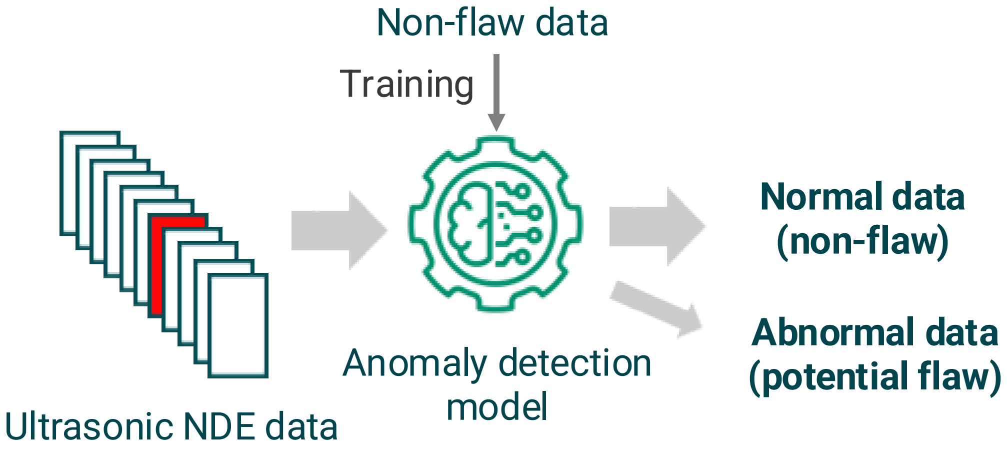
Near-term ML applications for ultrasonic NDE include assisted analysis and fully automated analysis



Automated Analysis problem formulation can change to accommodate dataset balance between flaws and non-flaws

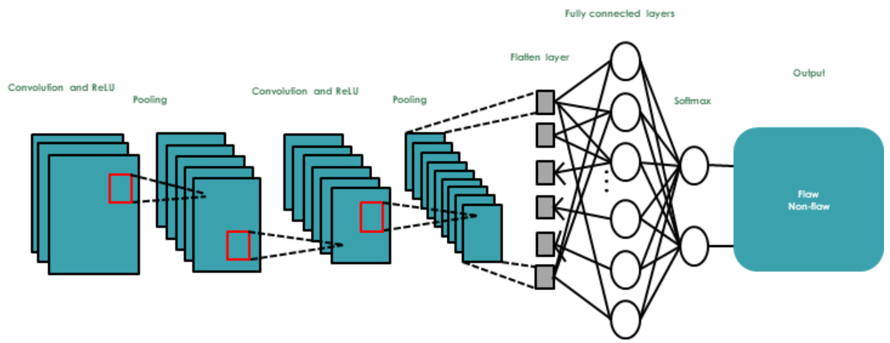


Automated Analysis: Classification

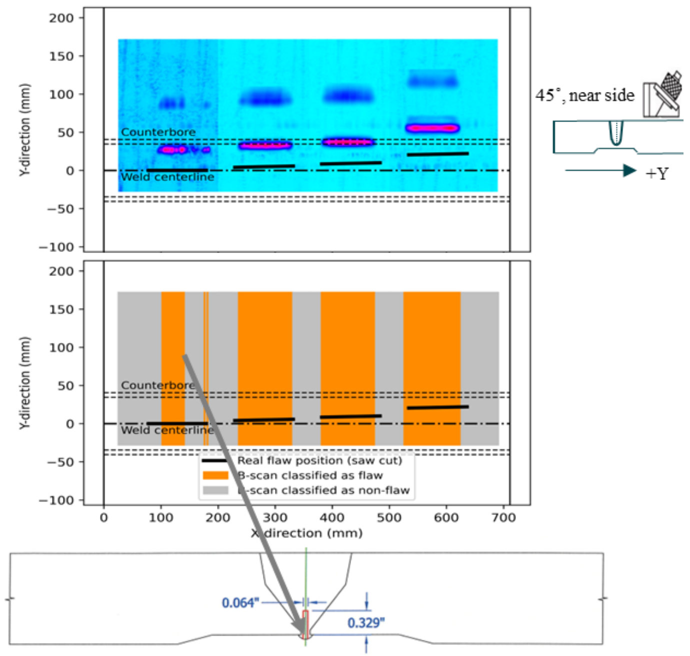


Automated Analysis: Screening

Recent results continue to show the potential of ML for automated NDE data analysis and…

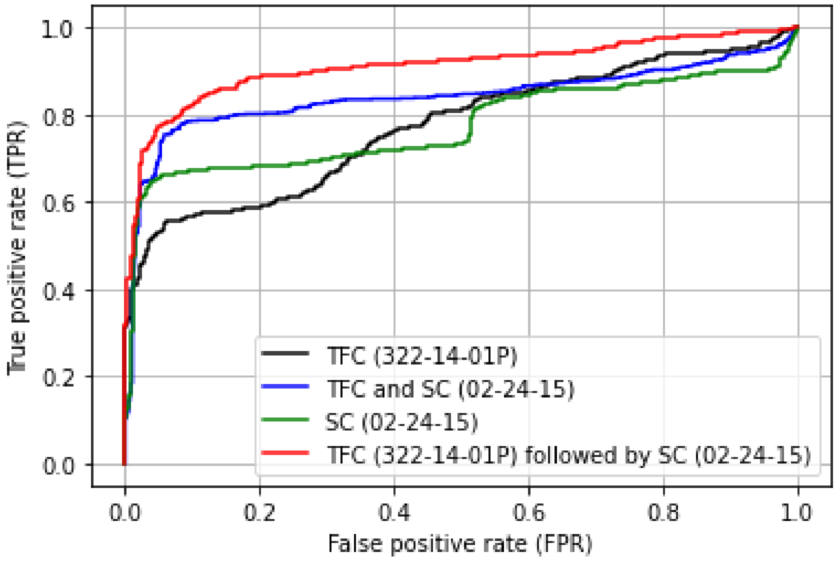


Example: Convolutional Neural Network Architecture



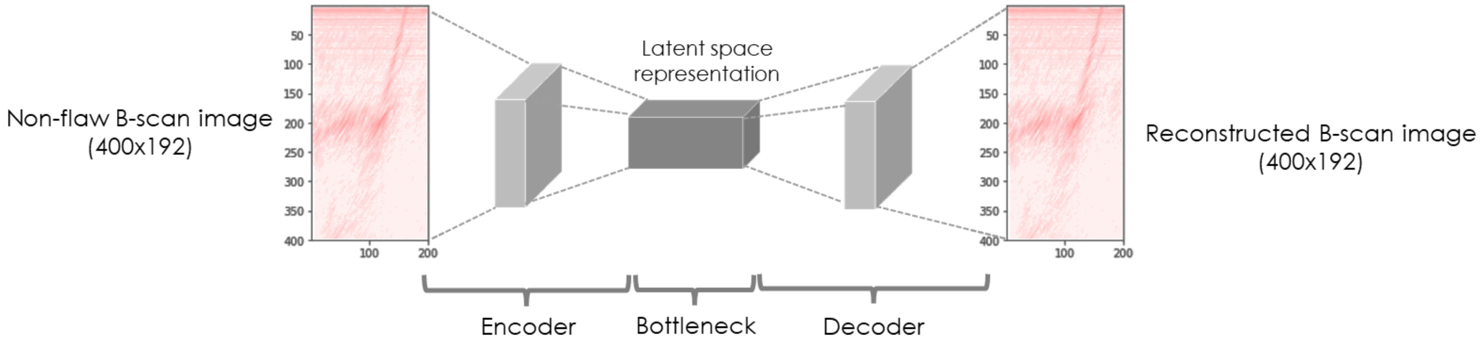
Flaw inside the weldment (TPR=0.54)

Example: Flaw Detection

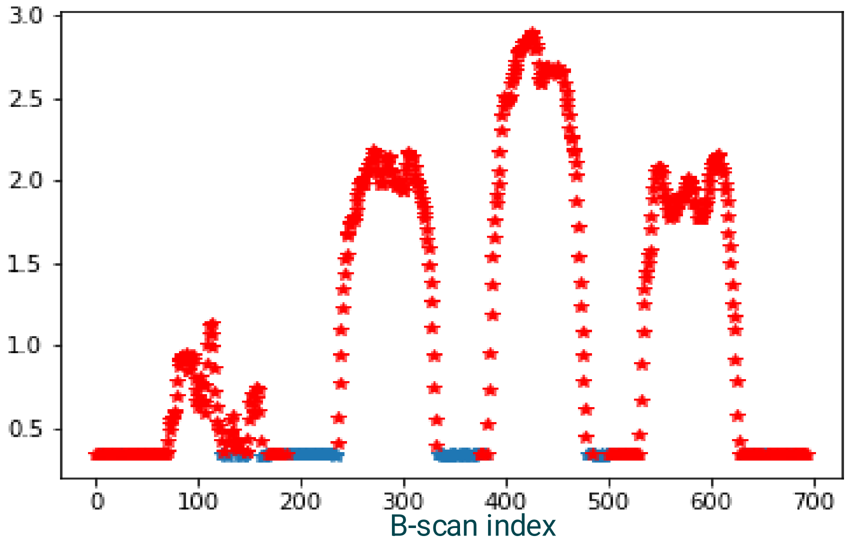


Receiver operating characteristic (ROC) curves

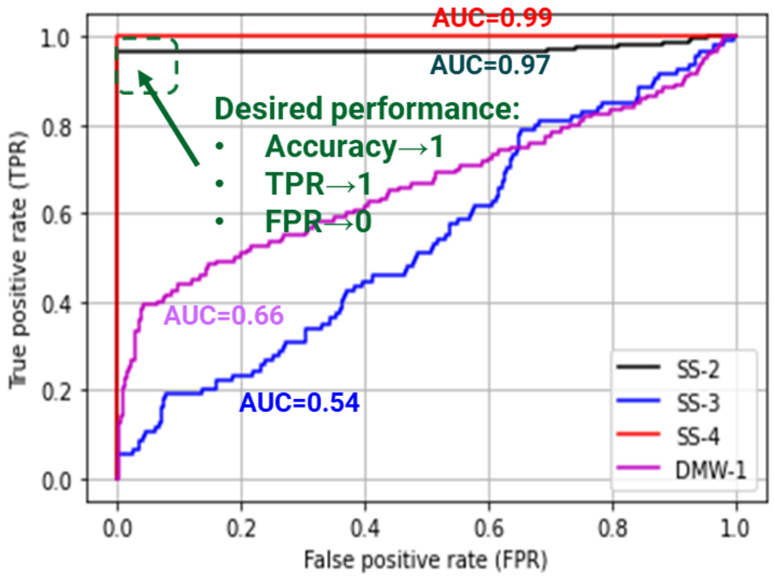
…ML for anomaly detection if applied appropriately



Example: Autoencoder Architecture



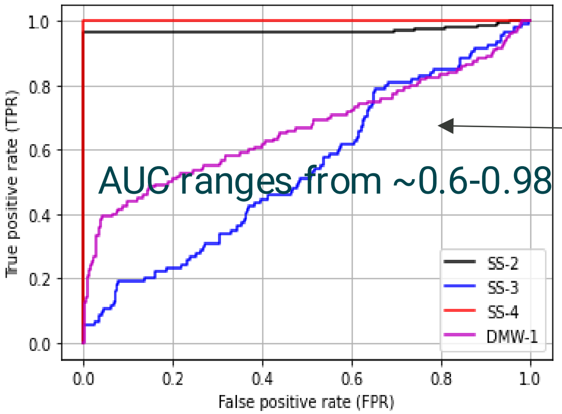
Example: Potential Flaw Identification



Receiver operating characteristic (ROC) curves and area under curve (AUC)

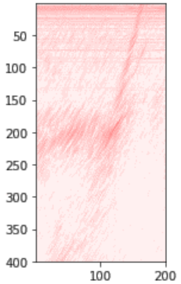
OAK RIDGE National Laboratory

Results indicate data diversity in the training data set is important for improving ML performance



AUC ranges from ~0.6-0.98

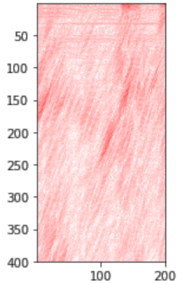
Anomaly Detection ROC Curves: before data augmentation



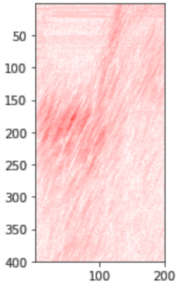
SS-1 (4 saw cuts)

Original Training Data Set

More Flaws Added



SS-3 (3 TFCs)

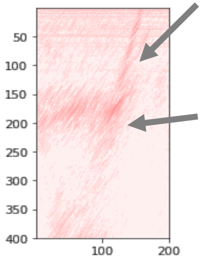


DMW-1 (4 TFCs)

Additional Data from other specimens

Augmented Data Added

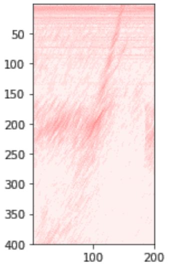
Window moved: up

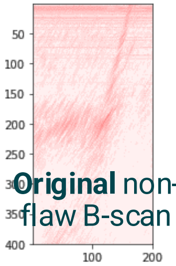


Right side of the weldment

Weld root

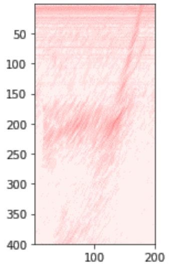
Window moved: **left**



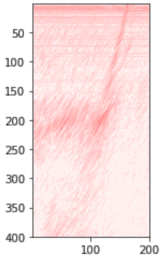


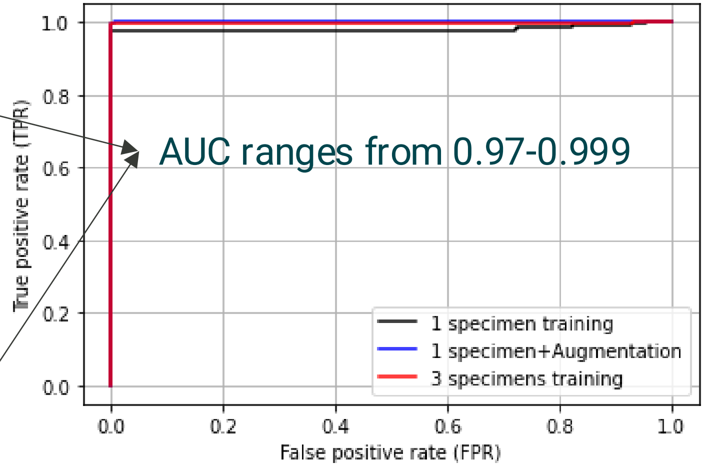
Original non-flaw B-scan

Window moved: **right**



1.25× amplitude





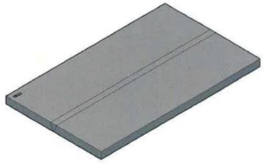
AUC ranges from 0.97-0.999

Anomaly Detection ROC Curves: after data augmentation

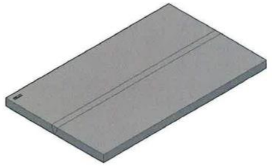
Reference NDE data set includes multiple probe designs, frequencies, and wave modes

A (19C-358-1)

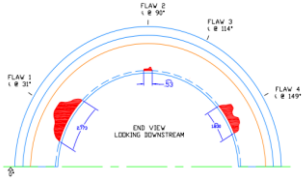
B (19C-358-2)



G (21C-303-1)

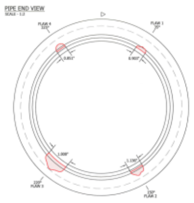


H (21C-303-3)

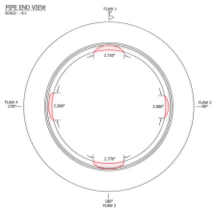


C (322-14-01P)

D (02-24-15)



E (8C-032)



F (8C-091)

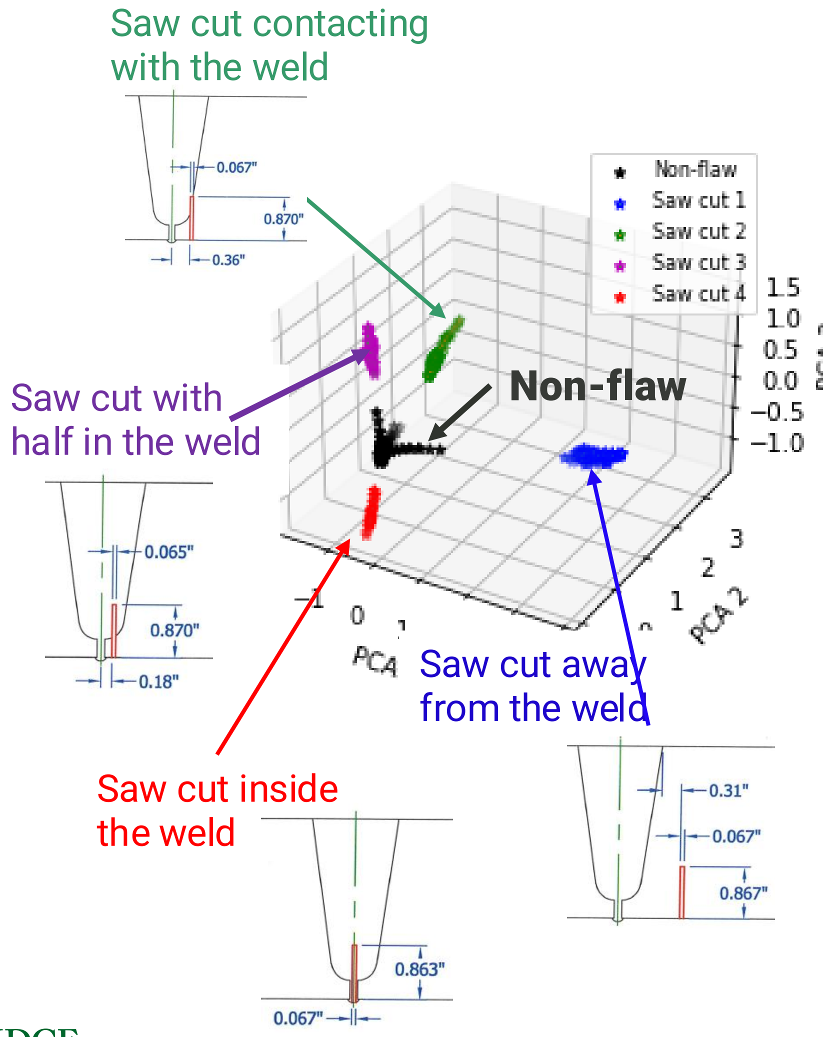
Specimens in Reference Data Set (To Date)

Specimen	Description	Flaw	Type	Flaw length (mm)	Height (% thickness
A(19C-358-1)	SS Plate	1	Saw cut	101.7	30.1%
		2	Saw cut	101.4	30.2%
		3	Saw cut	101.6	30.2%
		4	Saw cut	101.4	30.0%
B(19C-358-2)	SS Plate	1	Saw cut	100.6	29.2%
		2	Saw cut	101.4	29.2%
		3	Saw cut	101.4	29.4%
		4	Saw cut	101.4	29.5%
C(322-14-01P)	SS pipe section	1	TFC	70.4	65.8%
		2	TFC	13.5	12.5%
		3	TFC	46.5	43.0%
D(02-24-15)	SS pipe section	A	TFC	10.7	15.0%
		B	TFC	30.5	43.0%
		C	TFC	43.6	64.0%
		a	Saw cut	32.8	7.5%
		b	Saw cut	65.2	28.4%
		d	Saw cut	54.1	18.8%
		e	Saw cut	43.7	12.0%
E(8C-032)	DMW pipe	1	TFC	22.9	20.0%
		2	TFC	28.9	40.0%
		3	TFC	45.9	60.0%
		4	TFC	21.6	30.0%
F(8C-091)	DMW pipe	1	EDM notch	69.1	30.2%
		2	EDM notch	50.8	17.6%
		3	TFC	70.6	36.4%
		4	TFC	57.6	23.2%
G (21C-303-1)	SS plate	1	EDM notch	50.8	15.0%
		2	EDM notch	75.9	29.6%
		3	TFC	49.8	14.8%
		4	TFC	75.7	26.3%
H (21C-303-3)	SS plate	1	EDM notch	50.8	14.3%
		2	EDM notch	75.2	30.3%
		3	TFC	51.8	16.0%
		4	TFC	77.0	29.3%

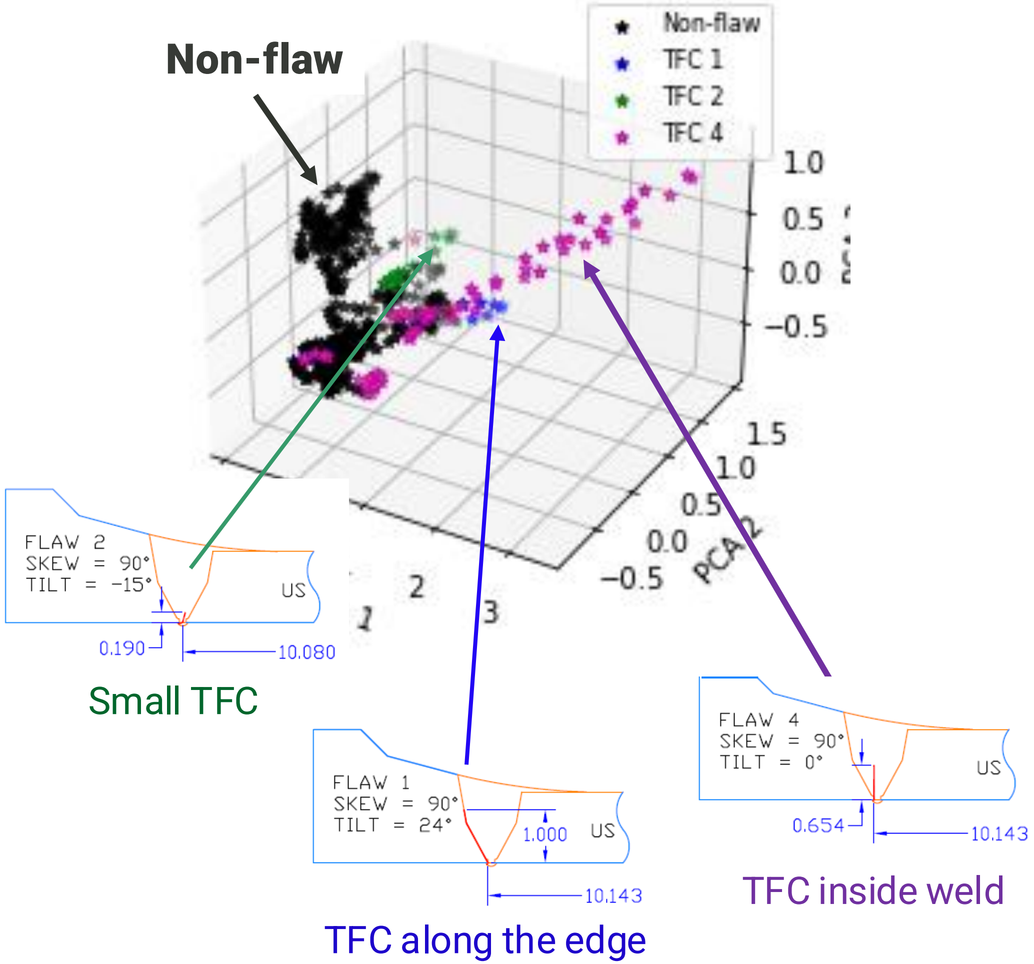
Flaws in Dataset (34 total: 12 saw cuts, 16 thermal fatigue cracks, 6 EDM notches)

Statistical data characterization points to differences in data from inspection parameters and specimen characteristics

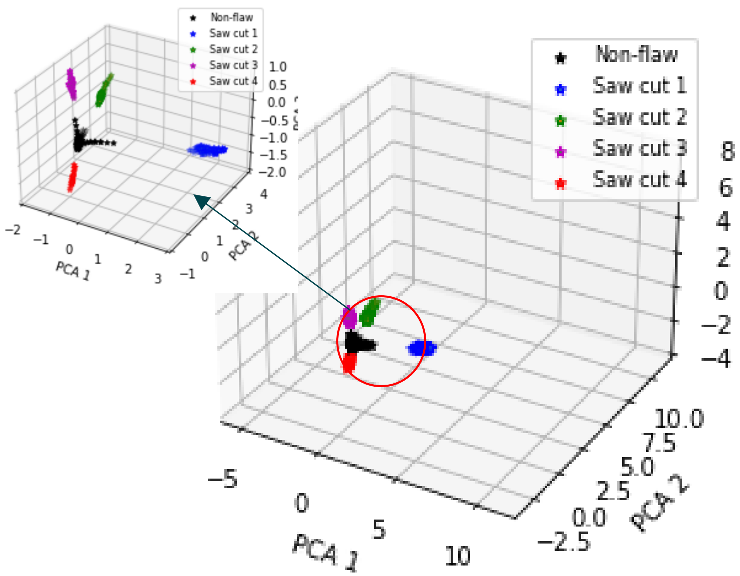
SS-1 (19C-358-1, 4 saw cuts)



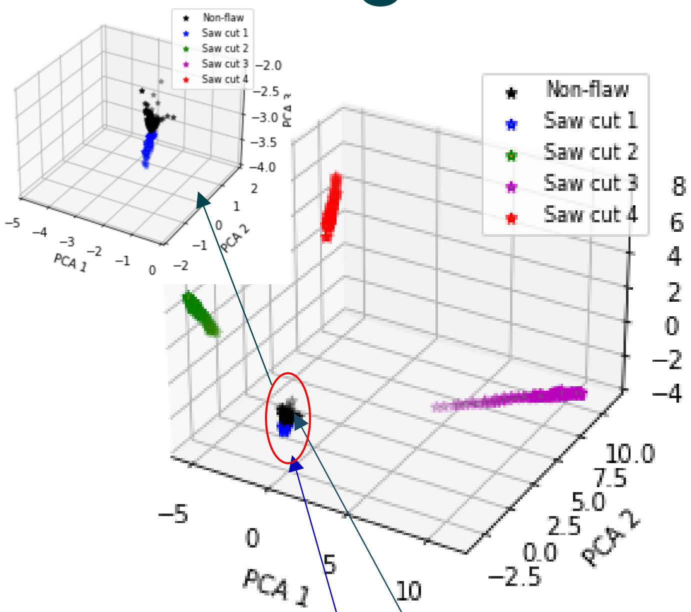
SS-3 (322-14-01P, 3 TFCs)



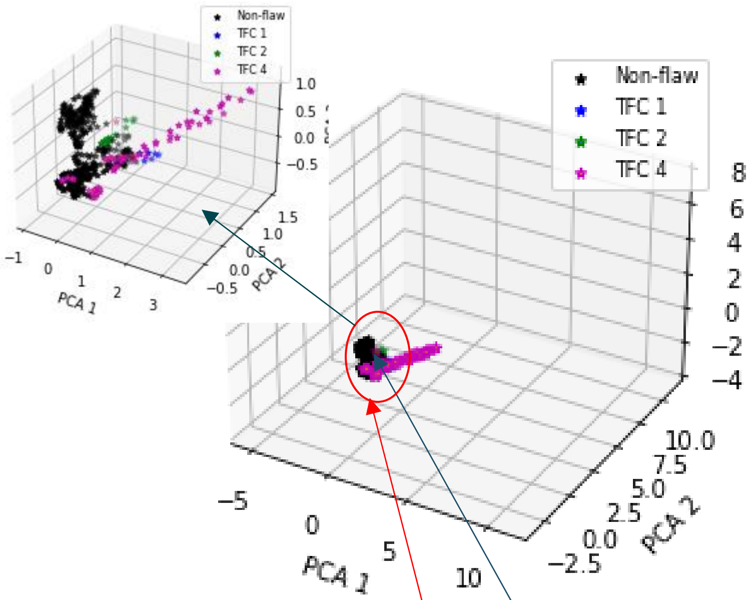
Statistical data characterization appears to highlight improved ML accuracy with similar data (single-element probe)



PCA, Specimen A (SwRI 45˚)



PCA Transformation from Specimen A, applied to Specimen B (SwRI 45˚)



PCA Transformation from Specimen A, applied to Specimen C (SwRI 45˚)

ML results (CNN model trained with Specimen A data)

		Actual value	
		Flaw	Non-flaw
Prediction	Flaw	298 (TP)	0 (FP)
	Non-flaw	29 (FN)	292 (TN)

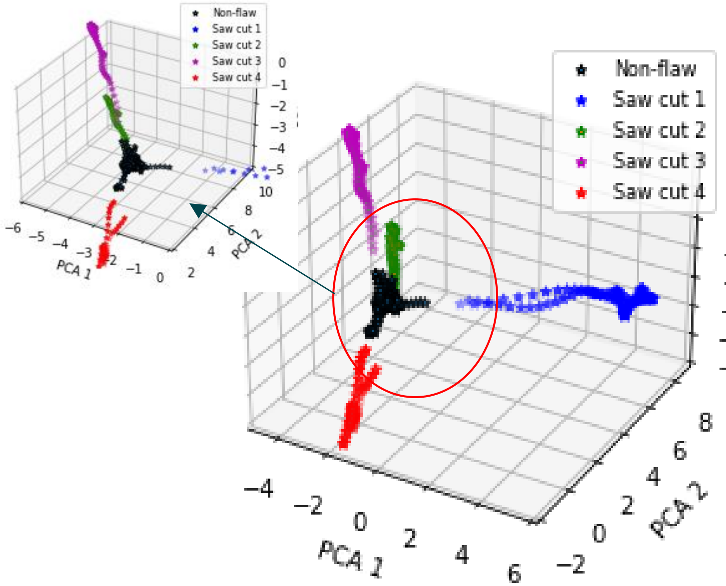
Accuracy=0.95, TPR=0.91, FPR=0

		Actual value	
		Flaw	Non-flaw
Prediction	Flaw	35 (TP)	8 (FP)
	Non-flaw	59 (FN)	378 (TN)

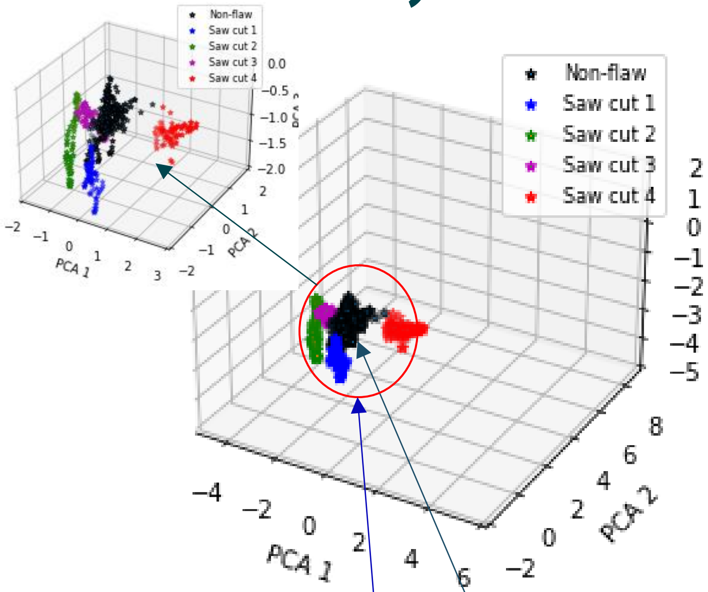
Accuracy=0.91, **TPR=0.62**, FPR=0.02

OAK RIDGE National Laboratory

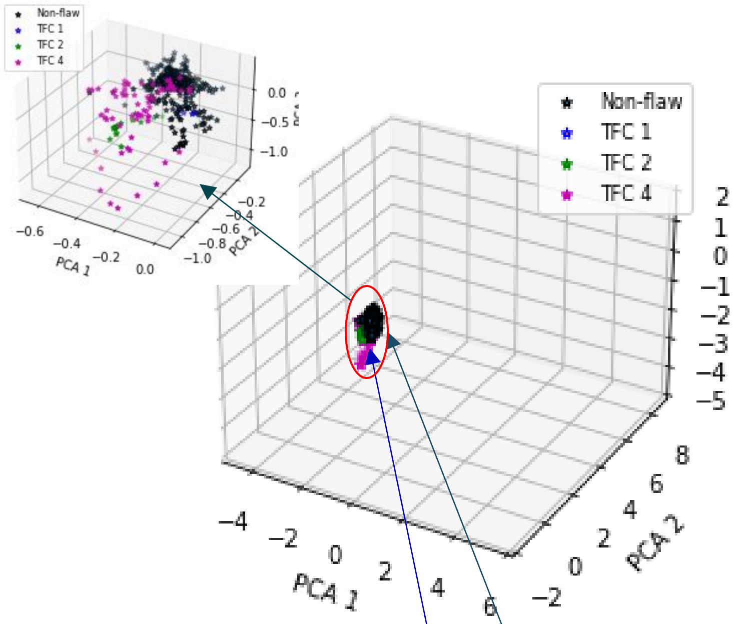
Statistical data characterization appears to highlight improved ML accuracy with similar data (PAUT)



PCA, Specimen A (PAUT 45˚)



PCA Transformation from Specimen A, applied to Specimen B (PAUT 45˚)



PCA Transformation from Specimen A, applied to Specimen C (PAUT 45˚)

ML results (CNN model trained with Specimen A data)

Actual value

Prediction	Flaw	Non-flaw
Flaw	301 (TP)	84 (FP)
Non-flaw	0 (FN)	309 (TN)

Accuracy=0.88,
TPR=1, FPR=0.2

Actual value

Prediction	Flaw	Non-flaw
Flaw	0 (TP)	0 (FP)
Non-flaw	128 (FN)	352 (TN)

Accuracy=0.73,
TPR=0, FPR=0

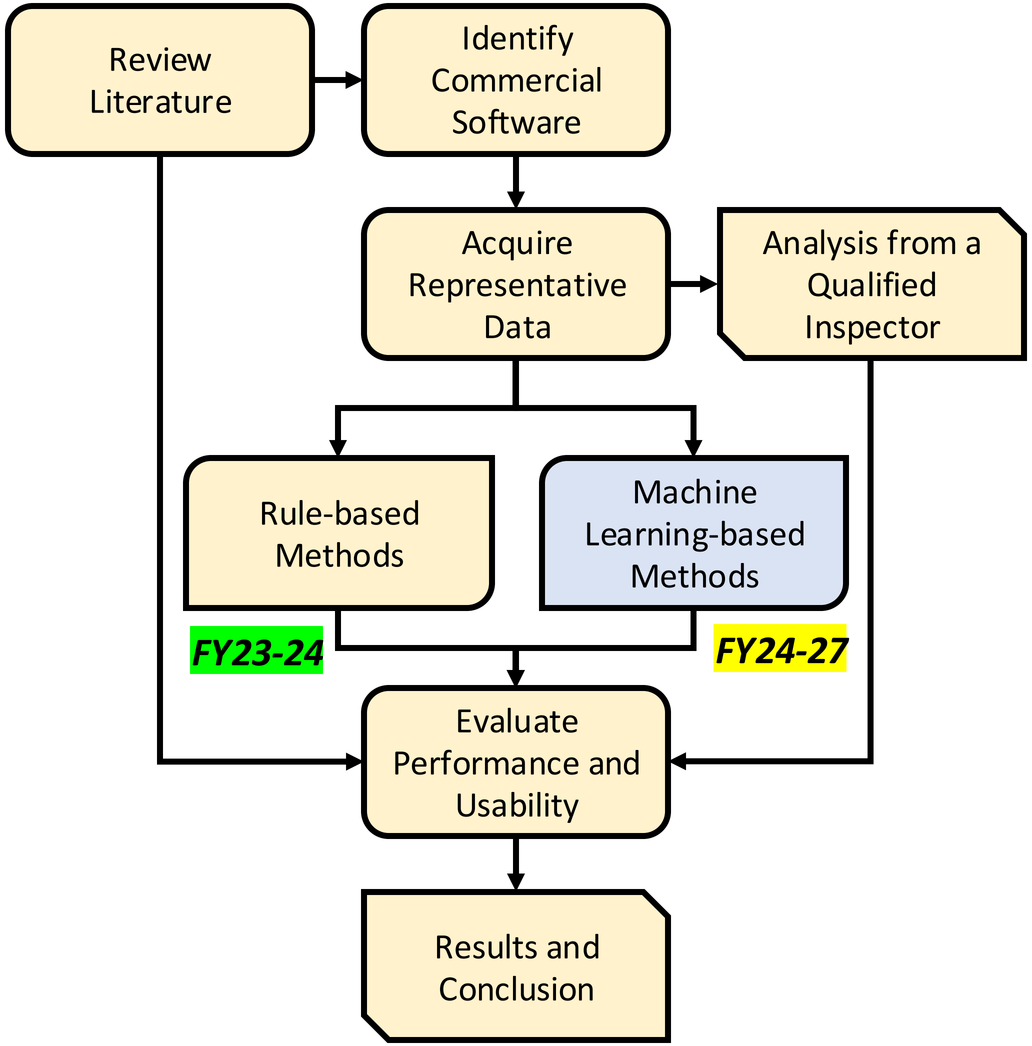
Oak Ridge National Laboratory

Planned work and expected outcomes

- Investigating ML and data verification and validation methods
- Evaluating data augmentation approaches and simulation data sets for training ML
- Examining model explainability approaches
- Metrics other than true positive and false positive rates for evaluating and monitoring ML performance

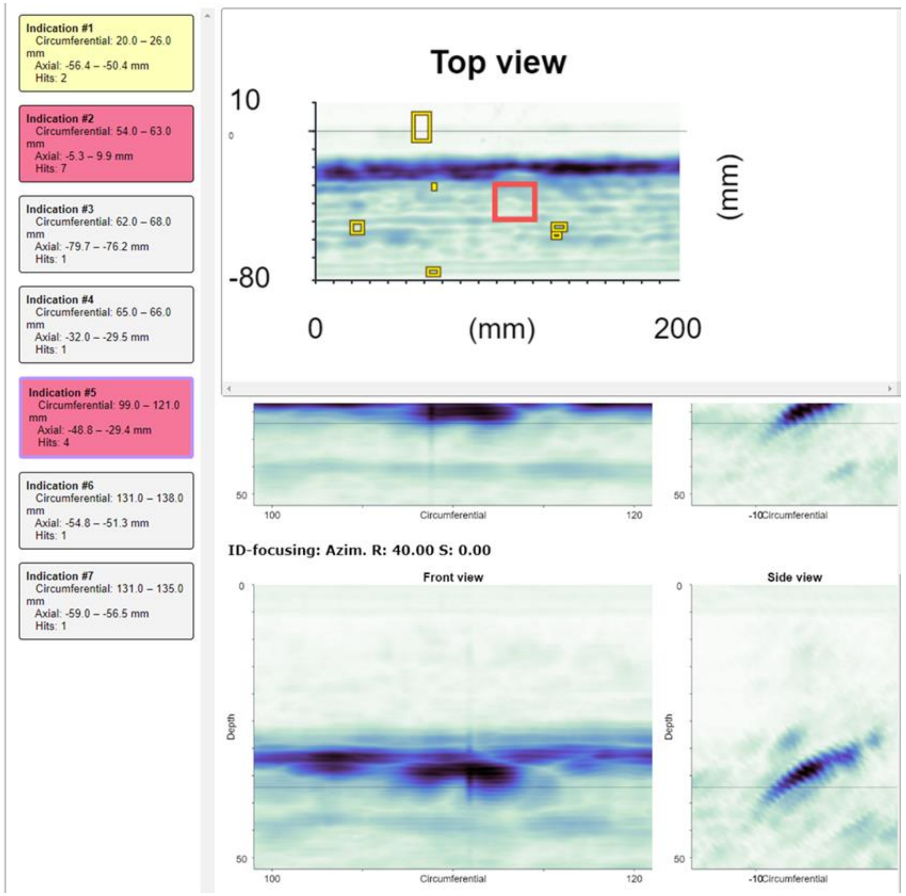
Confirmatory Analysis of Commercially Available Systems: Current Work

- PNNL is assessing commercial options of ML-based methods (E.g. TrueFlaw FlawML Unit, CIVA DataScience Module) for usability and capability.
- PNNL is performing an initial evaluation of TrueFlaw FlawML Edge (referred to as Box), using the DMW data collected on NRC-owned mockups available at PNNL.
- The DMW Model of FlawML Edge is trained and accepts data following EPRI-ENC-DMW-PA-1, EPRI's performance demonstration qualification summary and Supplement 10 inservice inspection (ISI) procedures for DMW examinations.
- PNNL has a wide array of available DMW mockups and access to Supplement 10. If necessary, PNNL will use these procedures to ensure acquired data are compatible with the FlawML Edge requirements.



Current Work (Continued)

- For each UVData file input into the DMW Model, FlawML Edge provides an HTML output file that displays the flaw indications, the number of hits for each indication, and the circumferential and axial axes of the indication.
  - If a UVData file is unable to process, FlawML Edge generates an error indicating “No Channels to Analyze”
- To analyze *.HTML output files, an evaluation table was developed to identify the following information
  - Mockup details
  - Flaw detected (Yes/No)
  - Number of hits
  - False positives (if any)
  - Number of hits on the false positive
  - Any other comments



Example of a FlawML Output Report - *.HTML File

Expected Outcomes and Planned Work

- Based on the data obtained from FlawML Edge unit, PNNL will evaluate the Machine Learning model through defined statistical metrics
  - ✓ Detection Rate—the total number of correctly identified flaws divided by the total number of flaws
  - ✓ Probability of Detection (POD)—the likelihood that flaws will be correctly identified within the inspection
  - ✓ False Call Probability (FCP)—the likelihood that flaws will be incorrectly identified within the inspection
  - ✓ Missed Detections—the number of flaws that were not identified within the inspection
- Future work will involve retraining of the ML models with TrueFlaw
- Further, an inspection of the CRDM model of FlawML Edge will be performed through NDE data acquired in CRDM mockups

Summary

- Focus on assessing current capabilities of machine learning (ML) and automated data analysis for improving NDE reliability
- Assessments utilizing empirical data sets
  - Available data covers austenitic and dissimilar metal welds
  - Multiple frequencies and probes
  - Generic procedures
  - Data cleaned and curated prior to use with ML
- Assessment of advances in ML, including commercially available ML systems, is ongoing

Questions?

Backup Slides

Codes and Standards for ML Remain a Work in Progress

- ML for NDE
  - ASME Boiler and Pressure Vessel Section V – Working Group on Automated Analysis
  - EPRI developing Code actions to address AI in ASME Boiler and Pressure Vessel Section XI, Appendix VIII
  - ASNT - best practices/guidance document development in process
- ML – definitions, trust, other applications, software requirements, software reliability, etc. are being addressed by other standards development organizations
  - ISO/IEC Joint Technical Committee (JTC) 1/SC 42 (Working group 42)
  - Institute of Electrical and Electronics Engineers (IEEE)
  - American Nuclear Society (ANS)
  - SAE

Current thinking seems to be on defining Codes and Standards requirements based on the criticality of the application of ML