



Introduction to mechanistic data-driven methods for engineering, mechanical science and mechanics of materials: difficulties in purely data-driven approaches for machine learning and some possible remedies

Prof. Wing Kam Liu, Walter P. Murphy Professor

Director of Global Center on Advanced Material Systems and Simulation (<https://camsim.northwestern.edu/>)

Northwestern University, w-liu@northwestern.edu

Students, Postdocs

Hengyang Li
Mahsa Tajdari
Satyajit Mojumder
Sourav Saha
Puikei Cheng
Orion L Kafka
Jiaying Gao
Cheng Yu
Dr. Zhengtao Gan

Collaborators

Greg Olson (NU, QuesTek)
NIST (Lyle Levine, Paul Witherell, Yan Lu)
CHiMaD (<https://chimad.northwestern.edu/>)
Jian Cao, Kori Ehmann
Greg Wagner
CT Wu, Zeliang Liu, and John Hallquist (LSTC)
Xianghui Xiao and Tao Sun (ANL)
John Sarwark (Feinberg NU, Lurie Hospital)



Outline

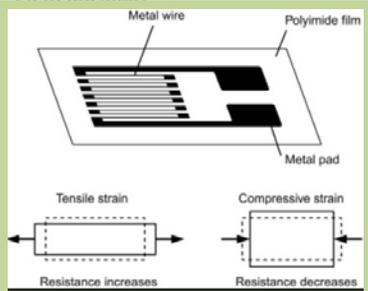
1. Motivation: sources of data in mechanical science and engineering
2. Mechanistic Machine Learning (MML) for mechanical science and engineering
 - Interpretation of the data
 - Relevant concepts in data science
 - Introduction to different Machine Learning (ML) methods
 - a. Unsupervised learning
 - b. Supervised learning
3. Applications of ML methods
 1. Topology optimization
 1. Feed Forward Neural Network (FFNN)
 2. FFNN+ Convolutional Neural Network (CNN)
 2. Adolescent Idiopathic Scoliosis
 1. FFNN
 2. Physics Guided Neural Network (PGNN)
4. Why we need reduced order models/methods (ROM)
5. Summary and conclusions
6. References



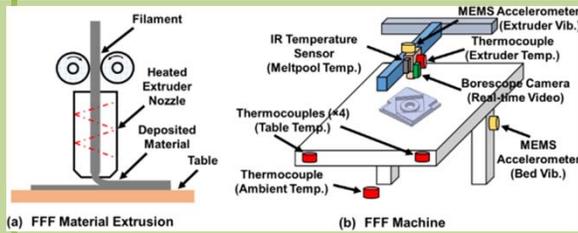
Outline

1. Motivation: source of data in mechanical science and engineering
2. Mechanistic Machine Learning (MML) for mechanical science and engineering
 - Interpretation of the data
 - Relevant concepts in data science
 - Introduction to different Machine Learning (ML) methods
 - a. Unsupervised learning
 - b. Supervised learning
3. Applications of ML methods
 1. Topology optimization
 1. Feed Forward Neural Network (FFNN)
 2. FFNN+ Convolutional Neural Network (CNN)
 2. Adolescent Idiopathic Scoliosis
 1. FFNN
 2. Physics Guided Neural Network (PGNN)
4. Why we need reduced order models/methods (ROM)
5. Summary and conclusions
6. References

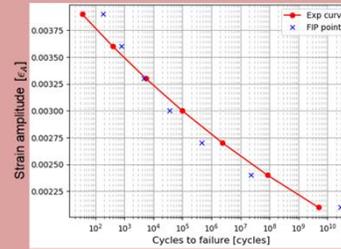
Multimodal data generation and collection



Strain Gauge^[1]



Temperature sensing in Additive Manufacturing (AM)^[2]



Fatigue life

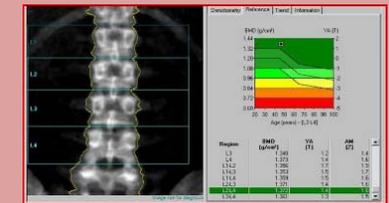


Courtesy Of NIST

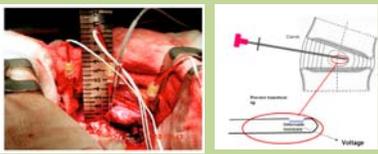
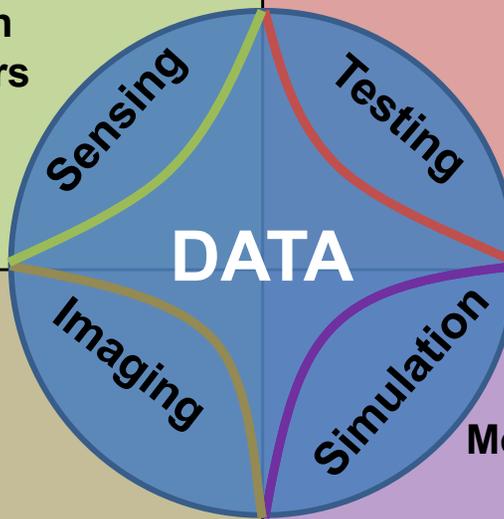
Composite coupon

Information from sensors

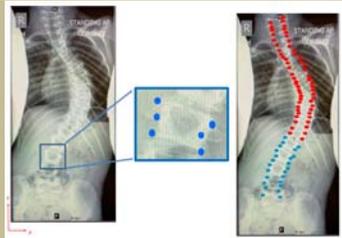
Test data



Measuring bone mineral density



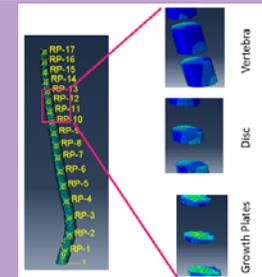
Pressure on intervertebral disc^[3]



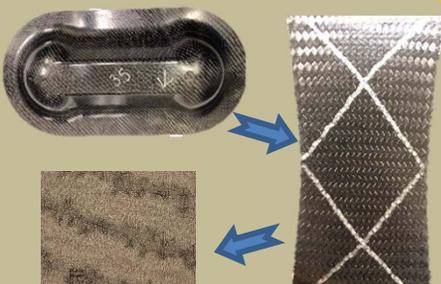
Extraction of data points from X-ray

2D/3D/4D images

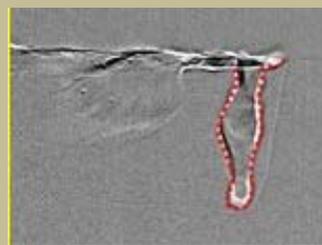
Model-based analysis



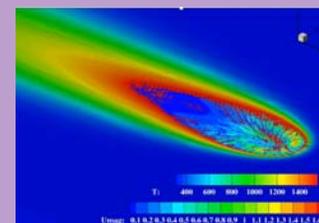
Spine simulation



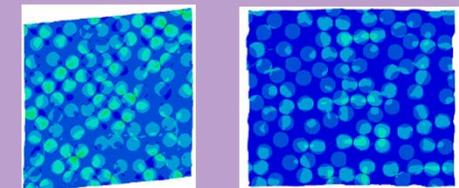
Composite structure



AM example



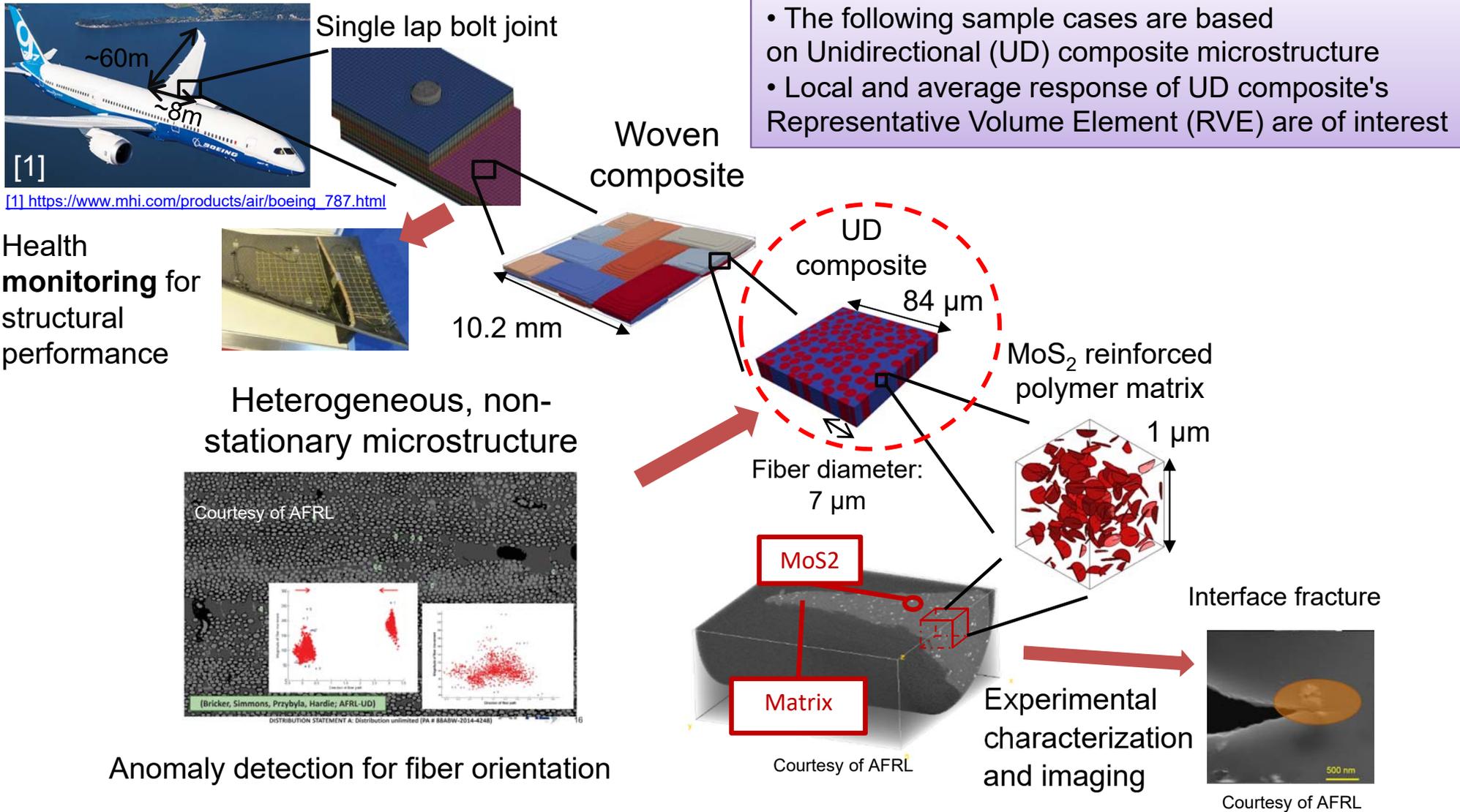
AM Process simulation



Composite materials simulation

Data generation and collection in composite systems

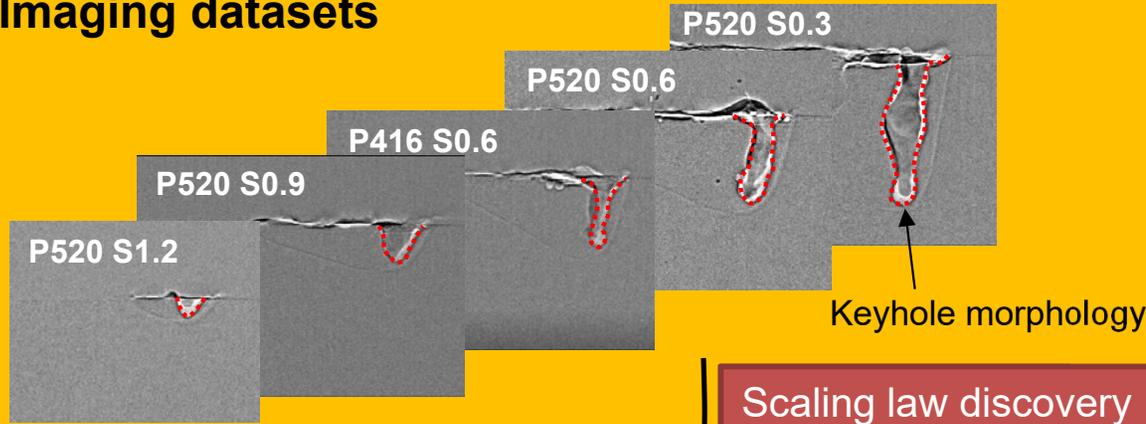
Data exist in multiple length scales for composite materials system



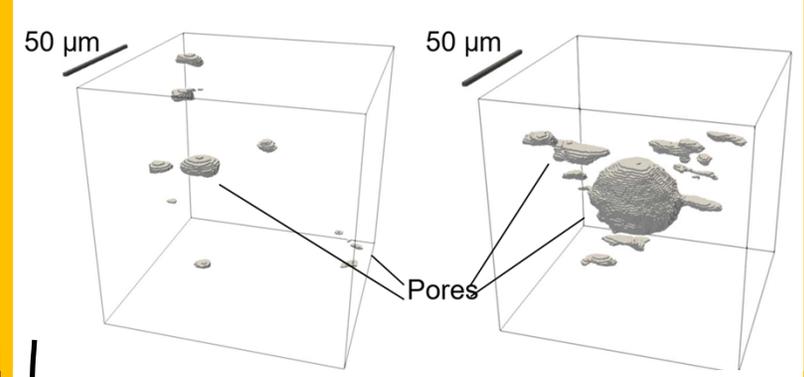
- Microstructure, material properties, structural performance.
- Information from four different scales are integrated to predict properties at part scale.

Approach: experiments and modeling motivated by data

Imaging datasets



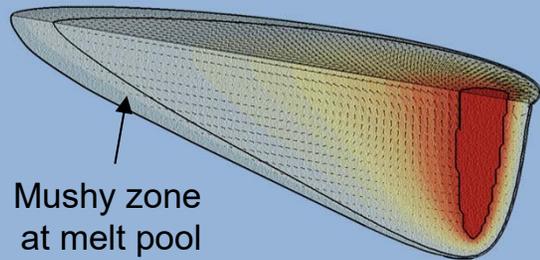
Porosity structure



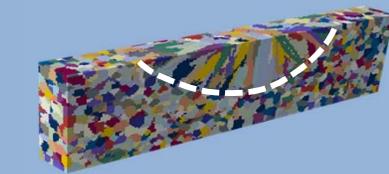
Scaling law discovery via machine learning

Image-based fatigue predictions

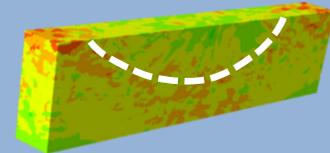
Data driven Process-structure-properties modeling



Mushy zone at melt pool boundary
(Cannot be observed by in-situ experiments)

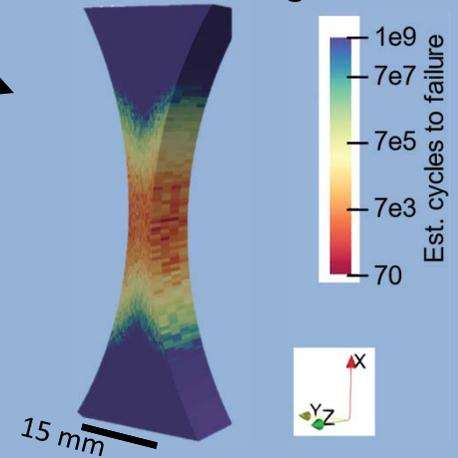


Cellular automaton (CA) modeling

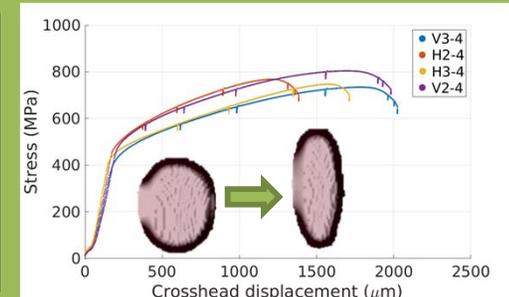
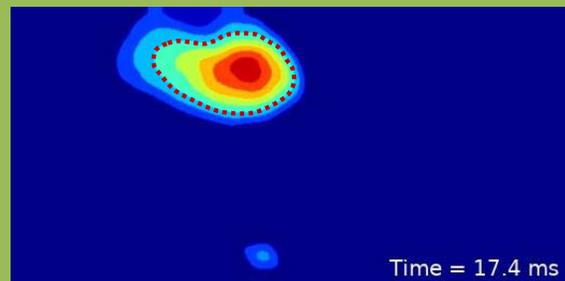
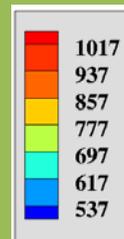
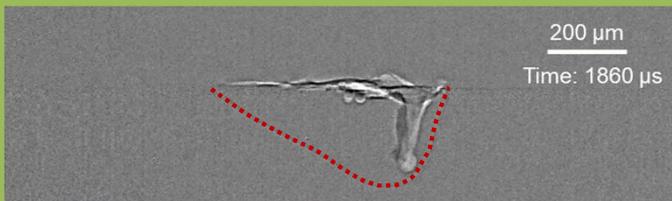


Grain-scale thermal-mechanical modeling

Fatigue life



Validation by experiments





Data generation and collection in Adolescent Idiopathic Scoliosis (AIS)

NORTHWESTERN UNIVERSITY

Spine images

*XR XR ** CT *** MR

Materialise Mimics

Multiple slices

Spinal angles (3D)

Thoracic Kyphosis Angle (α_1)

Trunk Inclination Angle (α_2)

Sacral Inclination Angle (α_4)

Lumbar Lordosis Angle (α_3)

Cobb Angle (α_5)

Sequential data

Age: t_1 months t_2 months t_3 months

All curvature types proposed by Scoliosis Research Society

Lumbar Thoracic

Double thoracolumbar

Bone mineral density

Vertebra	BMD (g/cm ³)	Z-score
C7	0.185	-1.2
T1	0.175	-1.5
T2	0.165	-1.8
T3	0.155	-2.1
T4	0.145	-2.4
T5	0.135	-2.7
T6	0.125	-3.0
T7	0.115	-3.3
T8	0.105	-3.6
T9	0.095	-3.9
T10	0.085	-4.2
T11	0.075	-4.5
T12	0.065	-4.8
L1	0.155	-2.1
L2	0.145	-2.4
L3	0.135	-2.7
L4	0.125	-3.0
L5	0.115	-3.3

Image segmentation (Assigning landmarks)

5 Lumbar 12 Thoracic

Post operated data

Spinal fusion

*XR: Xray ** CT: Computerized Tomography, ***MR: Magnetic Resonance, courtesy: Lurie Children's Hospital of Chicago



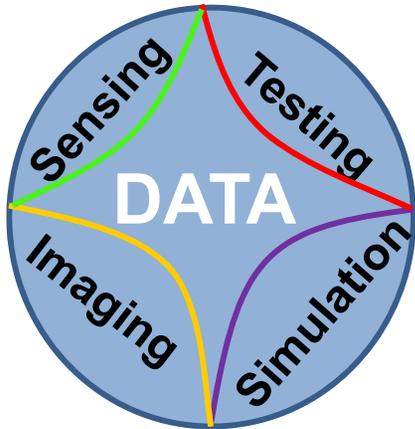
Outline

1. Motivation: source of data in mechanical science and engineering
2. Mechanistic Machine Learning (MML) for mechanical science and engineering
 - Interpretation of the data
 - Relevant concepts in data science
 - Introduction to different Machine Learning (ML) methods
 - a. Unsupervised learning
 - b. Supervised learning
3. Applications of ML methods
 1. Topology optimization
 1. Feed Forward Neural Network (FFNN)
 2. FFNN+ Convolutional Neural Network (CNN)
 2. Adolescent Idiopathic Scoliosis
 1. FFNN
 2. Physics Guided Neural Network (PGNN)
4. Why we need reduced order models/methods (ROM)
5. Summary and conclusions
6. References



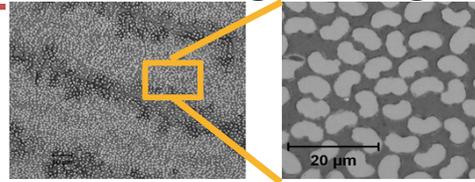
Interpretation of data in mechanical science and engineering

Data analysis using Machine Learning



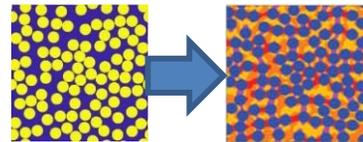
Means

Feature Engineering



Extraction of fiber dimension and fiber distribution from UD cross section

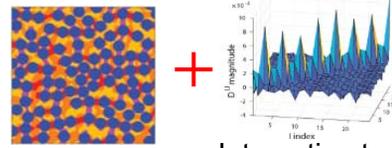
Dimension Reduction



600 x 600 voxels

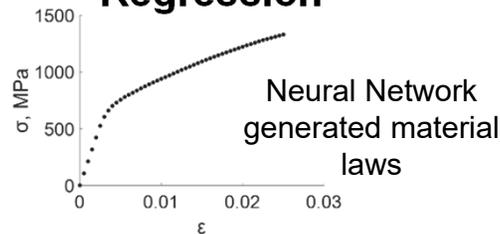
4 clusters

Reduced Order Models



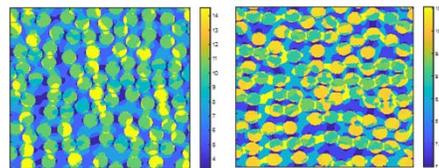
Interaction tensor

Regression



Neural Network generated material laws

Classification



Debonding

No debonding

Goals

To extract meaningful data

To reduce degrees of freedom

To speed up the computation

To discover hidden relationships

To categorize data





Outline

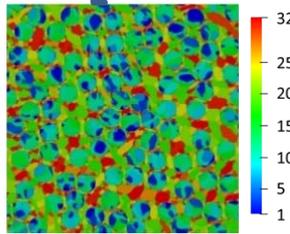
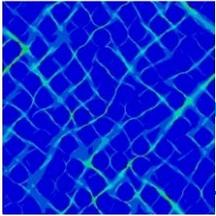
1. Motivation: source of data in mechanical science and engineering
2. Mechanistic Machine Learning (MML) for mechanical science and engineering
 - Interpretation of the data
 - **Relevant concepts in data science**
 - Introduction to different Machine Learning (ML) methods
 - a. Unsupervised learning
 - b. Supervised learning
3. Applications of ML methods
 1. Topology optimization
 1. Feed Forward Neural Network (FFNN)
 2. FFNN+ Convolutional Neural Network (CNN)
 2. Adolescent Idiopathic Scoliosis
 1. FFNN
 2. Physics Guided Neural Network (PGNN)
4. Why we need reduced order models/methods (ROM)
5. Summary and conclusions
6. References



Three types of machine learning in mechanical science and engineering

Unidirectional composite

Strain contour



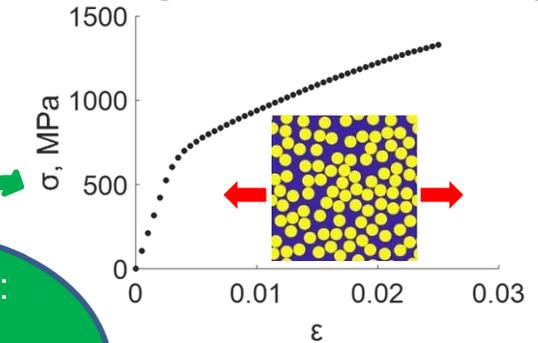
Clustering: grouping objects

Dimension reduction: reduces the number of features

e.g., Principal Component Analysis (PCA)

Unsupervised Learning: self-organized data pattern recognition

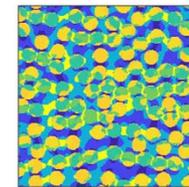
Predicts microstructure averaged stress given external loading



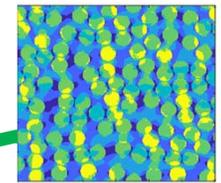
Machine Learning

Regression: Hidden relationship between variables

Damage detection by stress contour



No damage ✓



Damage ✗

Reinforcement Learning

Supervised Learning: mapping an input to an output

Classification: Identifying objects based on their class

Data from: Li, H., Kafka, O. L., Gao, J., Yu, C., Nie, Y., Zhang, L., Tajdari, M., Tang, S., Guo, X., Li, G., Tang, S., Cheng, G., & Liu, W. K. (2019). Clustering discretization methods for generation of material performance databases in machine learning and design optimization. Computational Mechanics, 1-25.



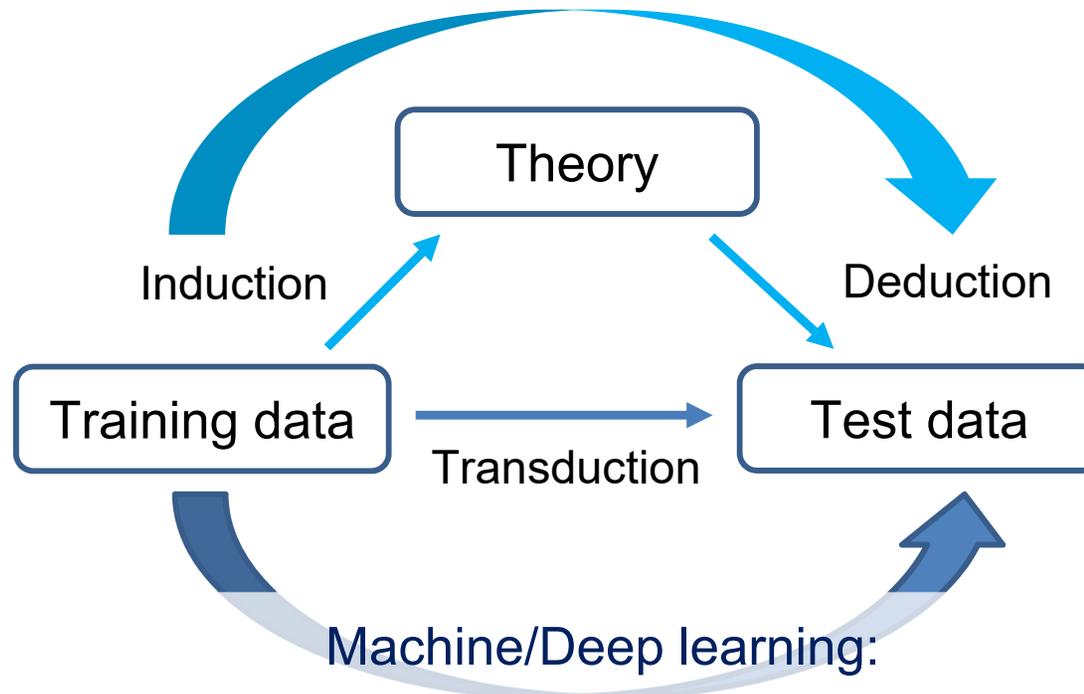
Data-Science: Transduction

Data-science is the “fourth paradigm” of science (empirical, theoretical, computational, data-driven) [21]

Conventional methods

Induction: specific observations to general theory (bottom-up)

Deduction: general theory to testable observations (top-down)



Transduction: learn from given data to apply to new data [22]



Relevant concepts in data science

Machine Learning

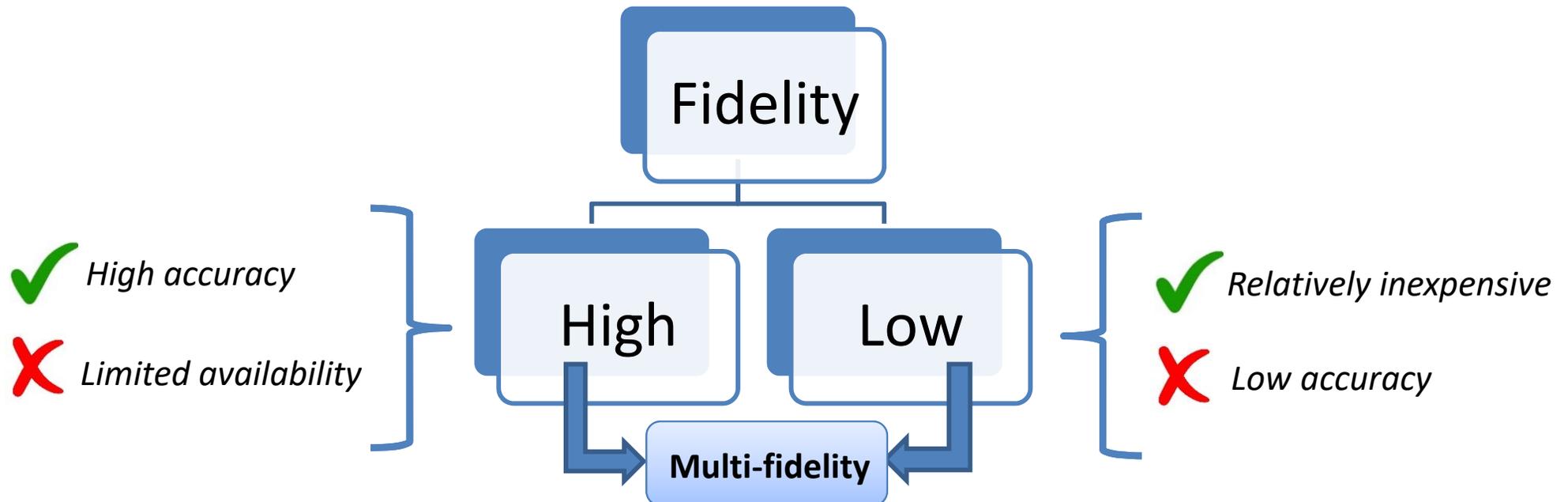
- ❑ A program or system that builds (trains) a predictive model from input data

Dimensionality

- ❑ Feature dimensionality: The number of features for each data point
- ❑ Input dimensionality: The total number of data points

Fidelity

- ❑ Quality of faithfulness of data





Relevant concepts in data science (cont)

Database

- A collection of rows or dataset with one or more features.

Features

- Individual independent variables defining characteristics of a data set.
- Informative and non-redundant data.

Feature engineering

- Process of determining which **features** might be useful and converting raw data into said features.

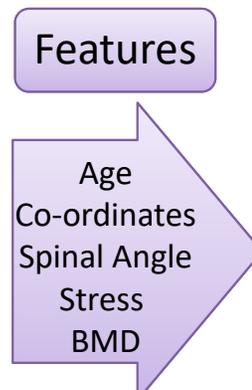
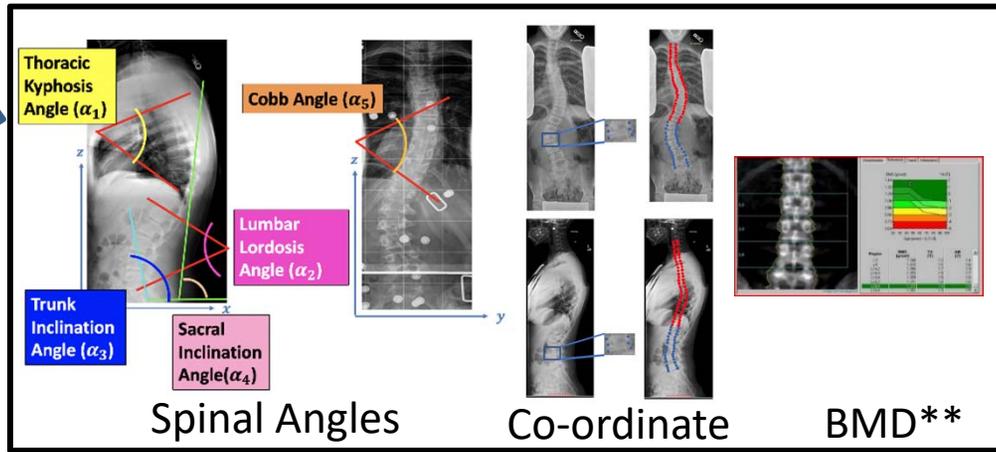
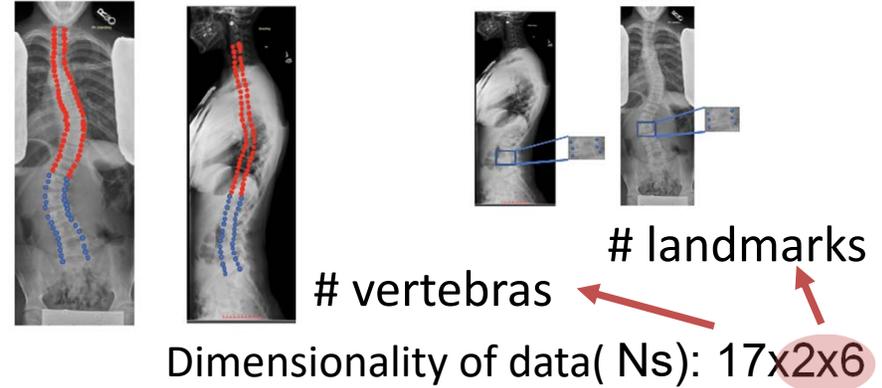
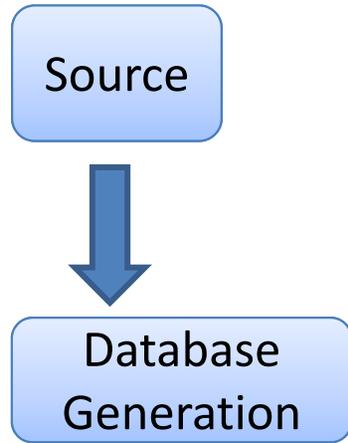
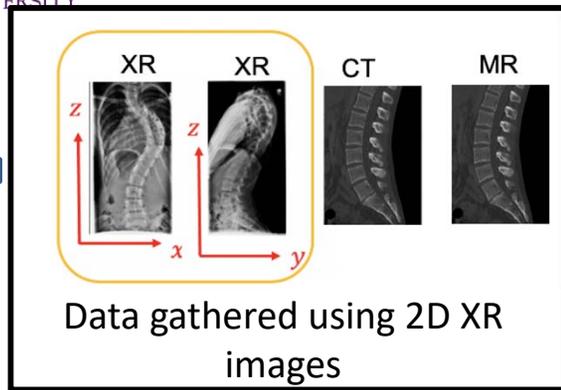
Dimension reduction

- Process of decreasing the number of dimensions representing a feature.

Objective function

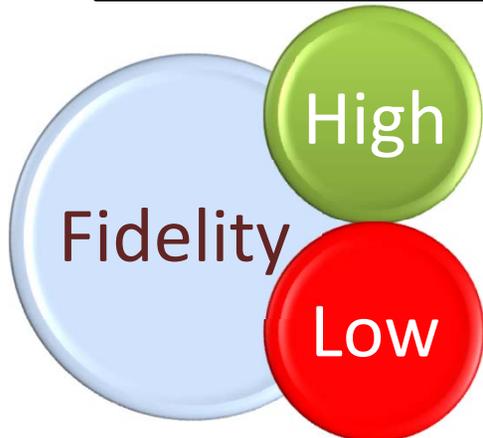
- The mathematical formula or metric that a model aims to optimize.

Illustration of relevant concepts for AIS*



Dimensionality of features

Features	Data points					
	1	2	3	.	.	Ns
X						
σ
α
t
Δt						
BMD						



- Co-ordinates from x-ray
- Age, test frequency
- Spinal angles
- Stress

X = Vector of input coordinates of a landmark $[X_1 \ X_2 \ X_3]$

σ = Stress vector $[\sigma_{11} \ \sigma_{22} \ \sigma_{33} \ \sigma_{12} \ \sigma_{23} \ \sigma_{31}]$

α = Global angle (Spinal Angles) vector $[\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4 \ \alpha_5]$

t = Age of the patient

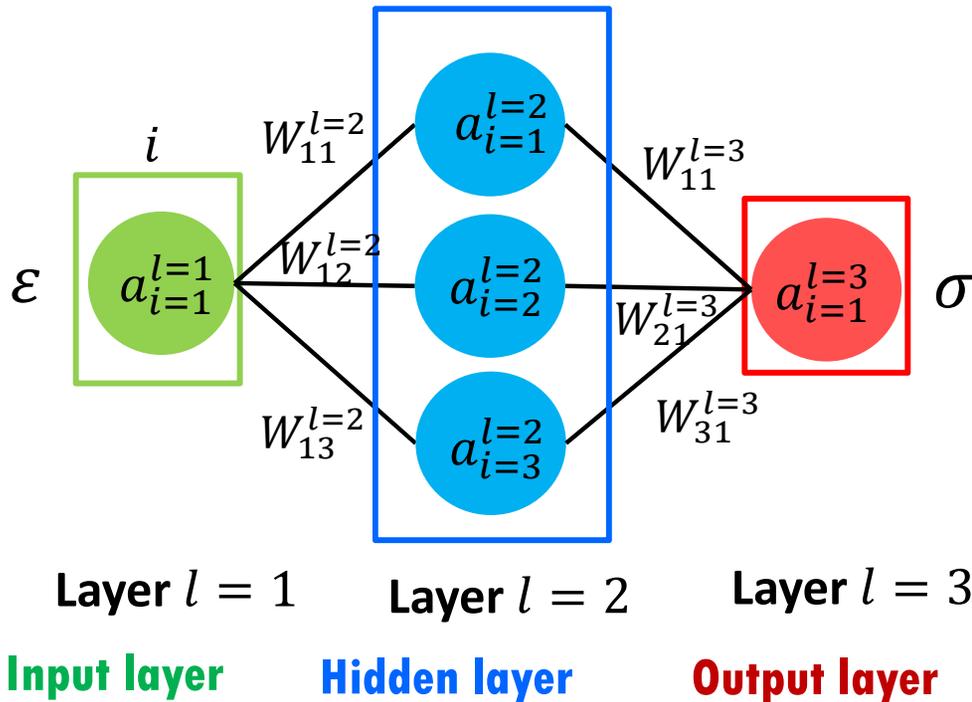
Δt = age variance between target age and current age (month)

*Adolescent Idiopathic Scoliosis **Bone Mineral Density



Basic concepts of artificial neural network (ANN)

Objective: To learn hidden relationship between input and output



Sample input data

	ϵ	σ (MPa)
Data point 1	0.1	20
Data point 2	0.2	38.6
...		

Assume $\sigma^* = 20$

Gradient descent^[1] :

$$\Delta W_{11}^{l=3} = \alpha \delta a_{i=1}^{l=2} \quad \Delta W_{11}^{l=2} = \alpha \delta W_{11}^{l=3} a_{i=1}^{l=1}$$

$$\Delta W_{21}^{l=3} = \alpha \delta a_{i=2}^{l=2} \quad \Delta W_{12}^{l=2} = \alpha \delta W_{12}^{l=3} a_{i=1}^{l=1}$$

$$\Delta W_{31}^{l=3} = \alpha \delta a_{i=3}^{l=2} \quad \Delta W_{13}^{l=2} = \alpha \delta W_{13}^{l=3} a_{i=1}^{l=1}$$

σ^* : target value α : learning rate

$$\delta = (\sigma^* - a_{i=1}^{l=3})$$

Optimization problem:

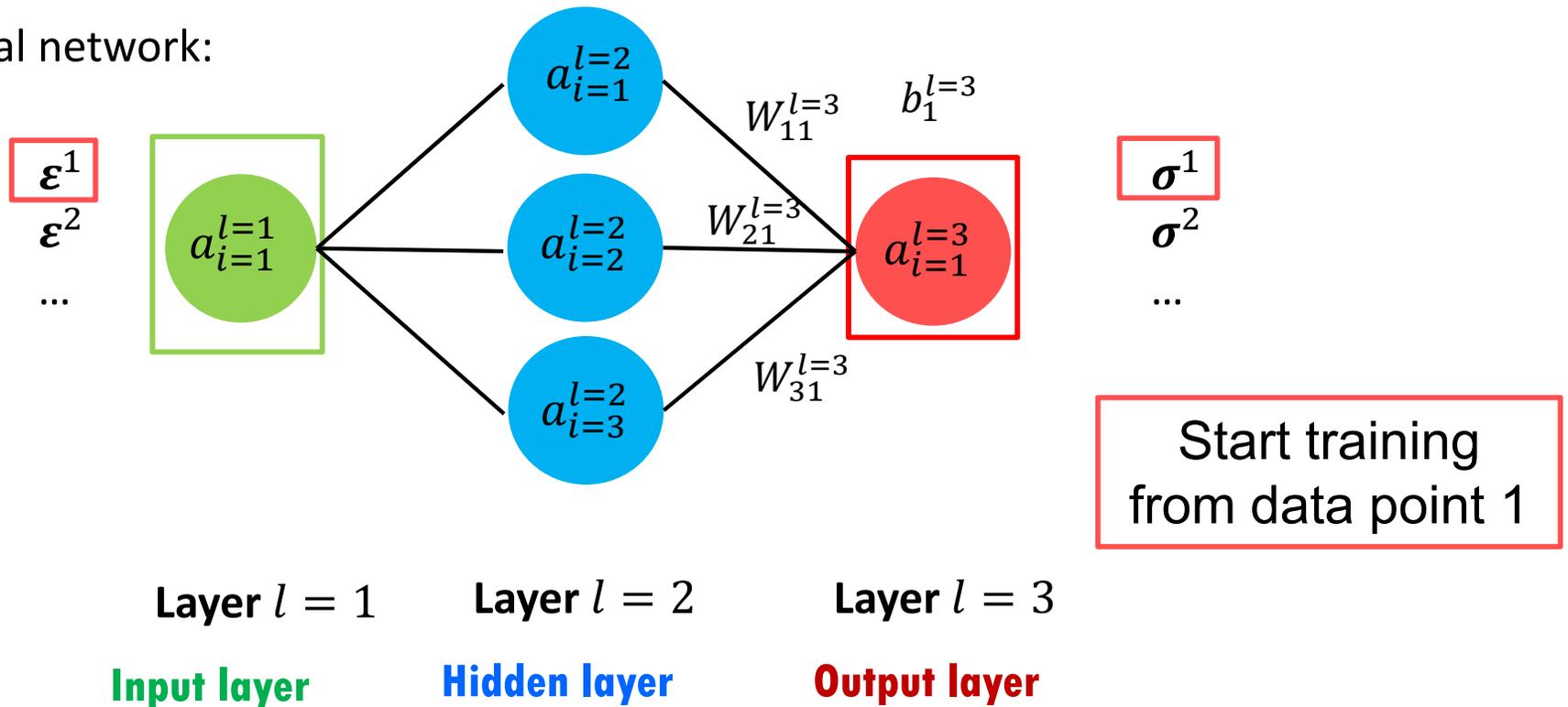
$$\text{minimize Error: } E = \frac{1}{2} (\sigma^* - a_{i=1}^{l=3})^2$$

[1] Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.



Example of training Neural Network (NN): learning back-propagation

Neural network:



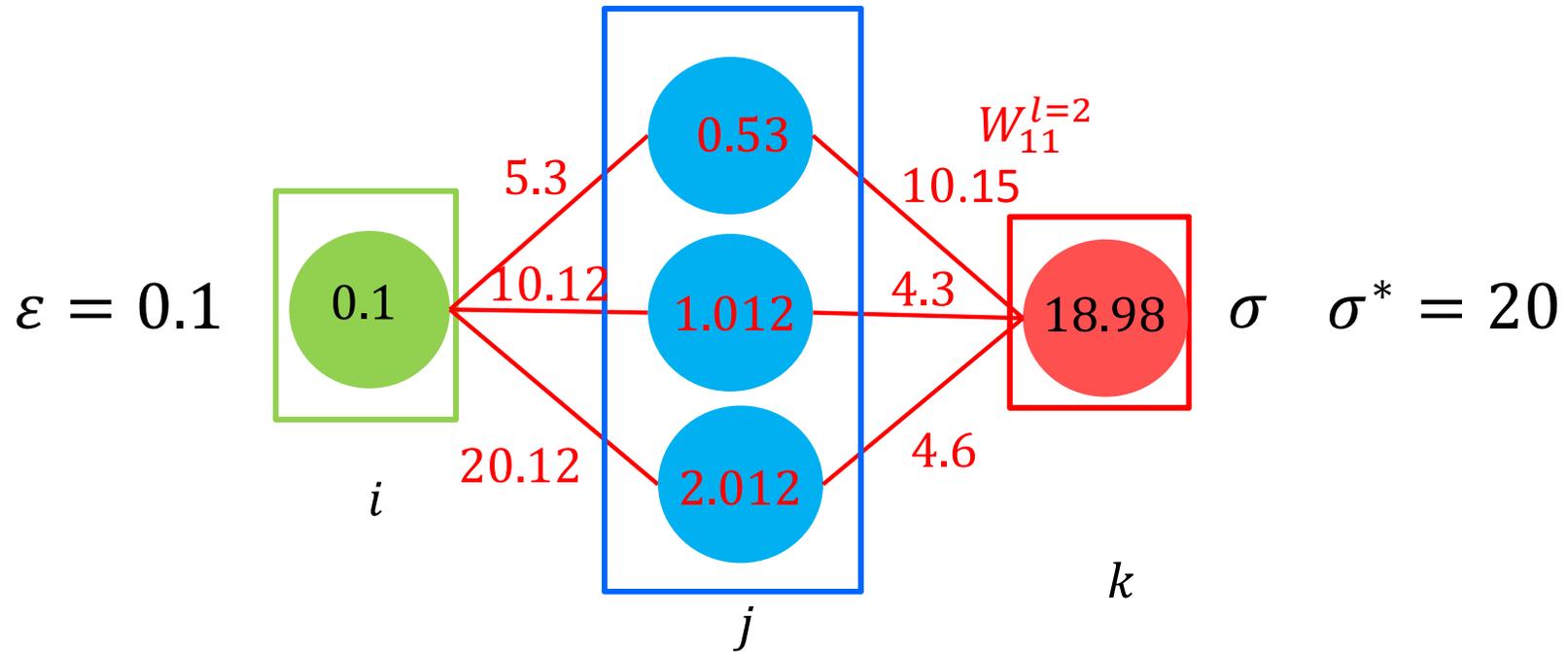
Start training from data point 1

$$a_1^{l=3} = f\left(\sum_{i=1}^3 W_{i,1}^{l=3} a_i^{l=2} + b_1^{l=3}\right)$$

$f(X)$: activation function,
 e.g. ReLU function, $f(X) = \max(0, X)$



Check error, and iterate for convergence



$$\text{Error: } E = \frac{1}{2} (\sigma^* - a_{k=1}^{l=3})^2$$

$$E = 4.5 \rightarrow 0.5202$$

The error will reduce by iteration, finally

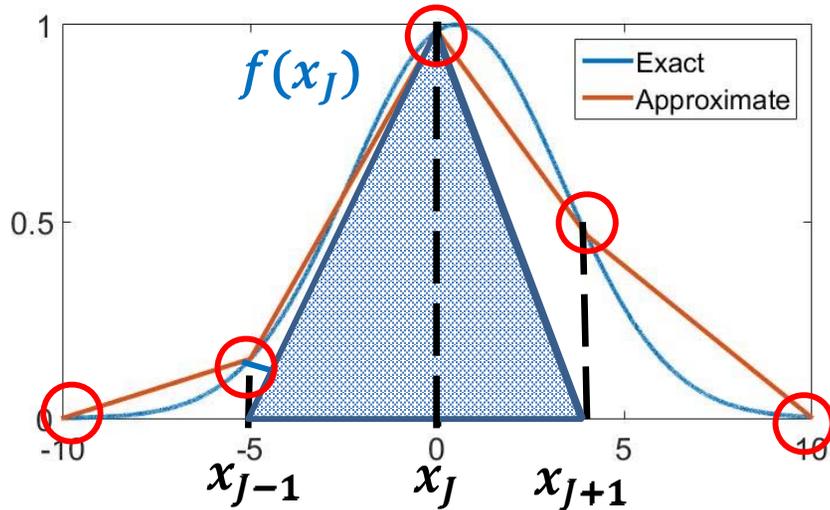
$$E \leq E^*, \text{ convergence}$$

Repeat for all data points until error is minimized



Neural Network as interpolation function^[1]

We often use a linear interpolation function $f^h(x)$ to approximate any continuous function $f(x)$.



- ◆ Goal function $f(x)$:
e.g. $f(x) = \exp(-(x - 0.5)^2/16)$ (blue line)

- ◆ Approximate function: N_f Nodes

e.g. $\sum_{J=1}^{N_f=5} f(x_J)N(x; x_J)$,

$\{x_1, x_2, \dots, x_5\} = \{-10, -5, 0, 4, 10\}$
(Orange line)

- ◆ Assume linear shape function

ReLU: Rectified linear unit, usually in form: $\max(0, X)$

$$\sum_{J=1}^{N_f=5} f(x_J)N(x; x_J) = \sum_{J=1}^{N_f=5} \underbrace{T_J[x; \text{ReLU}, x_J, x_{J-1}, x_{J+1}]}_{\text{Shape function}} f(x_J)$$

Shape function approximated by NN^[2, 3] $\Leftrightarrow G^N(x; \{\mathbf{w}\}_J, \{\mathbf{b}\}_J)$

[1] Zhang, L., Yang, Y., Li H., Gao J., Reno D., Tang S., Liu W.K. Neural network finite element method, in preparation

[2] [Approximation by superpositions of a sigmoidal function](#), by George Cybenko (1989).

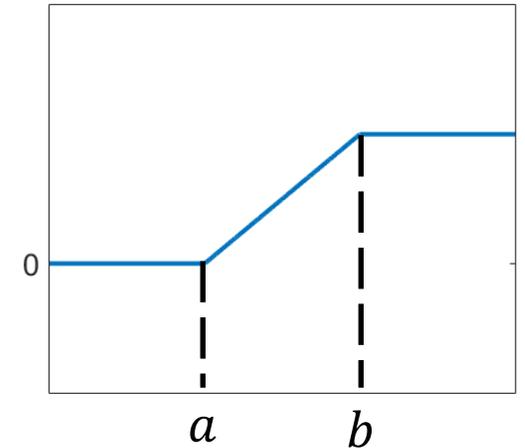
[3] [Multilayer feedforward networks are universal approximators](#), by Kurt Hornik, Maxwell Stinchcombe, and Halbert White (1989).



Proof: NN for 1D shape function approximation

Lemma 1 The continuous piece-wise linear function

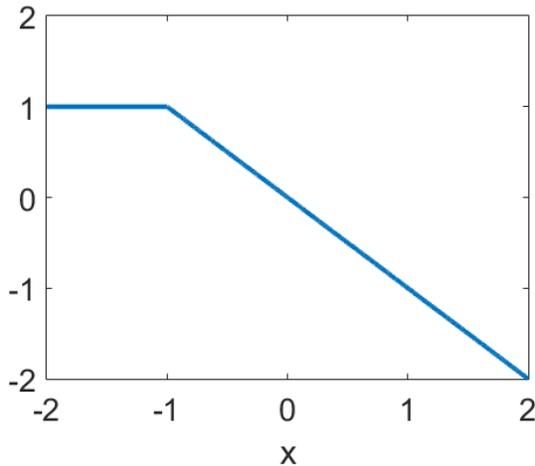
$$B(x; a, b) = \begin{cases} 0 & x < a \\ x - a & a \leq x \leq b \\ b - a & x > b \end{cases},$$



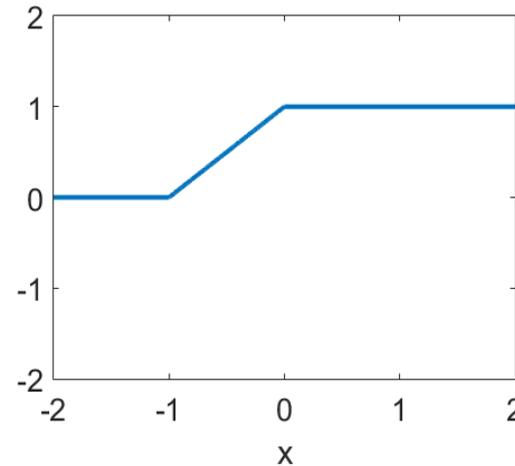
can be represented by neural network as

$$-ReLU(-ReLU(x - a) + b - a) + b - a$$

e.g. $a=-1, b=0$



Step 1: $-ReLU(x + 1) + 1$

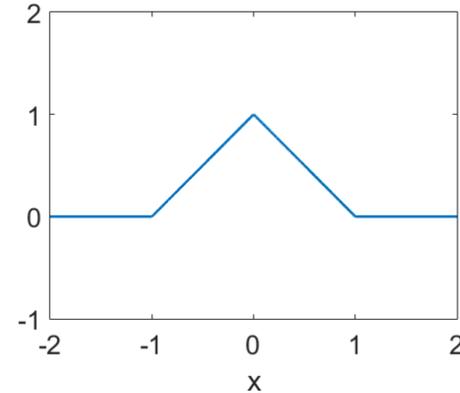
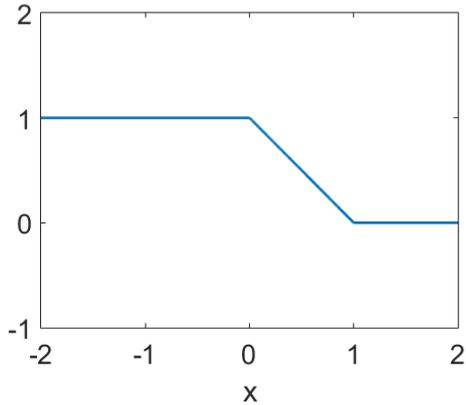


Step 2: $-ReLU(-ReLU(x + 1) + 1) + 1$



Proof: NN for 1D shape function approximation

For 1D linear basis function, take the reflection to construct the right part and then combine these two parts.

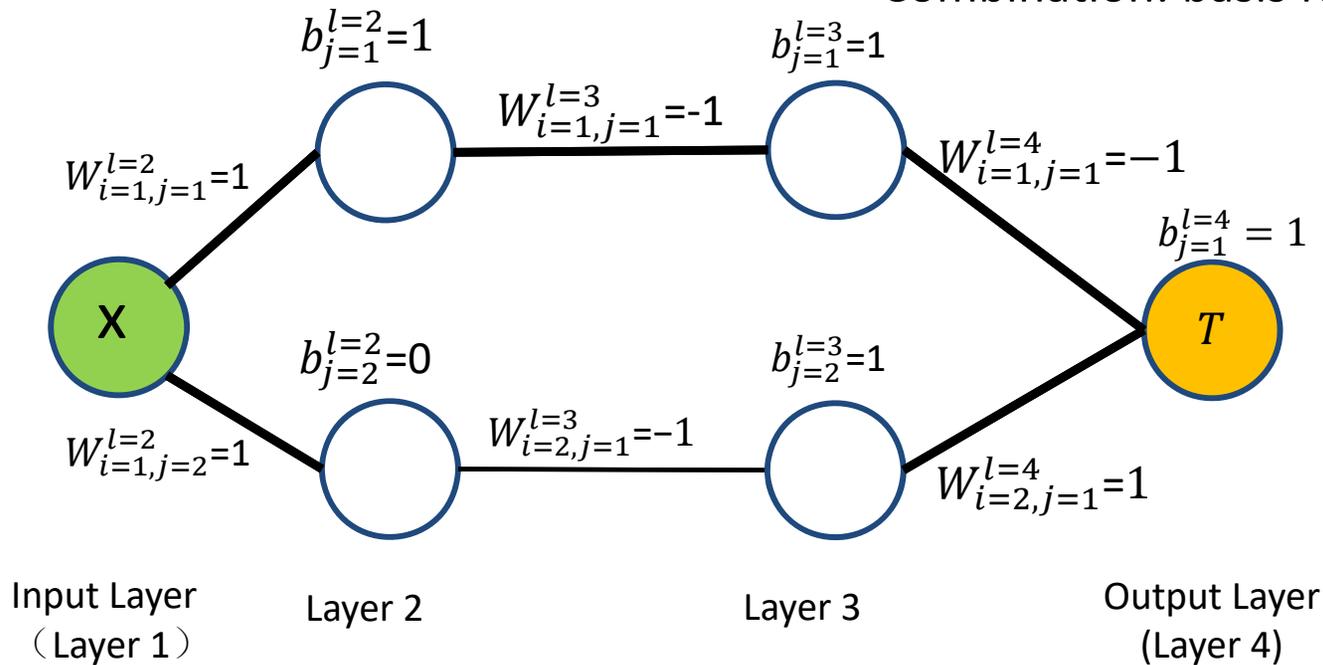


Step 3: $ReLU(-ReLU(x) + 1)$

Step 4: $-ReLU(-ReLU(x + 1) + 1) + ReLU(-ReLU(x) + 1)$

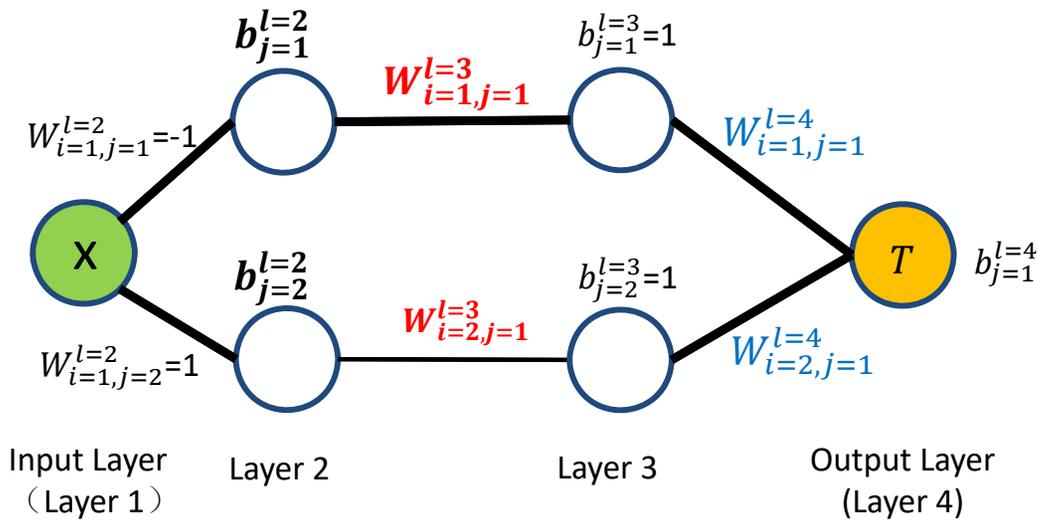
Reflection: right part

Combination: basis function

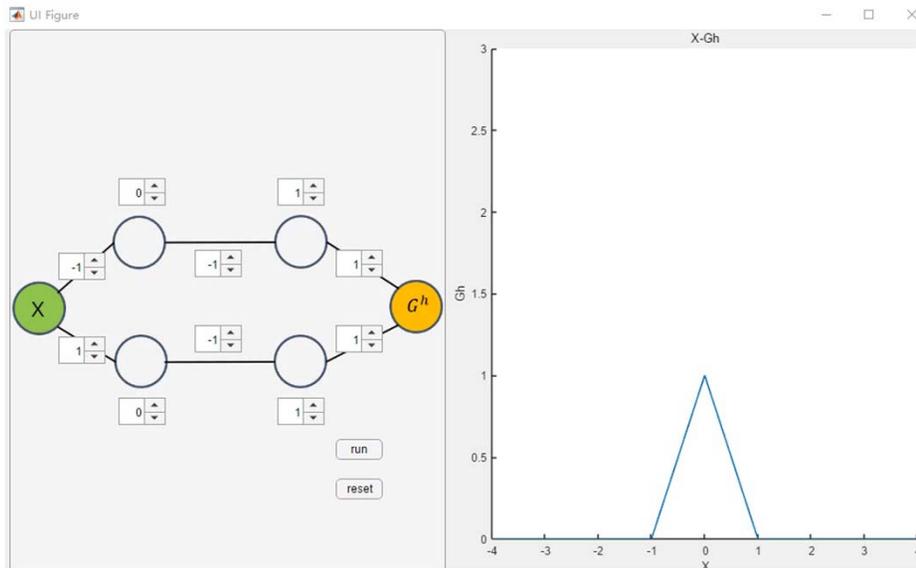




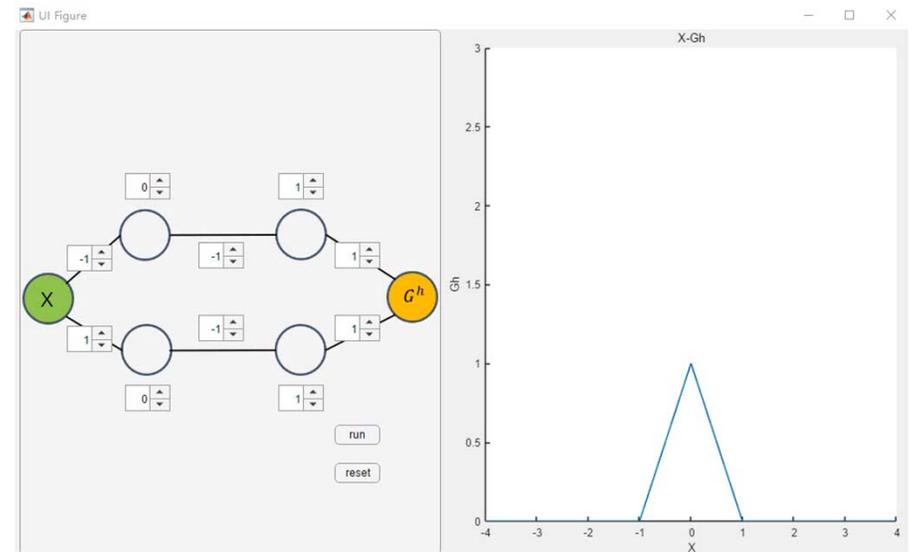
Effect of weights and biases on the output



	X	T
Data point 1	1	0
Data point 2	0.5	0.3
...		



Change in bias $b_j^{l=2}$, changes **the location**



Change in weights $W_{i=1,j=1}^{l=3}$, changes **the slope**



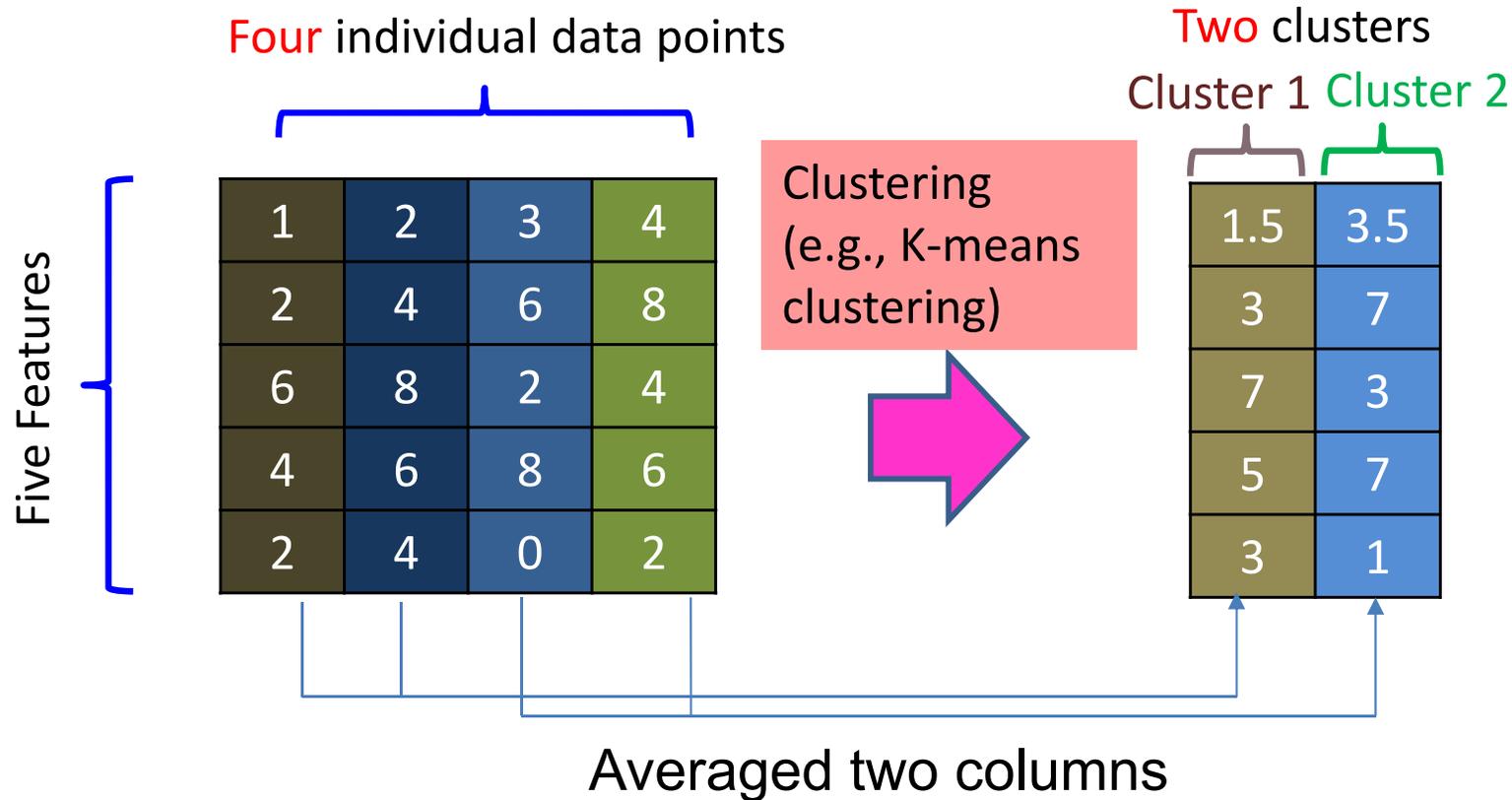
Outline

1. Motivation: source of data in mechanical science and engineering
2. Mechanistic Machine Learning (MML) for mechanical science and engineering
 - Interpretation of the data
 - Relevant concepts in data science
 - Introduction to different Machine Learning (ML) methods
 - a. Unsupervised learning
 - b. Supervised learning
3. Applications of ML methods
 1. Topology optimization
 1. Feed Forward Neural Network (FFNN)
 2. FFNN+ Convolutional Neural Network (CNN)
 2. Adolescent Idiopathic Scoliosis
 1. FFNN
 2. Physics Guided Neural Network (PGNN)
4. Why we need reduced order models/methods (ROM)
5. Summary and conclusions
6. References



A simple illustration on unsupervised learning for clustering

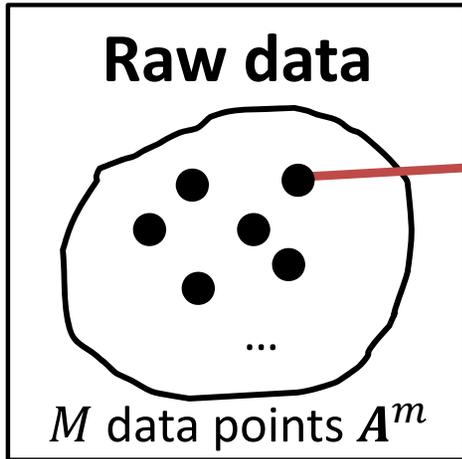
Objective: Group 4 data points (each having five features) into 2 clusters



Clustering: Reduces the data dimensionality

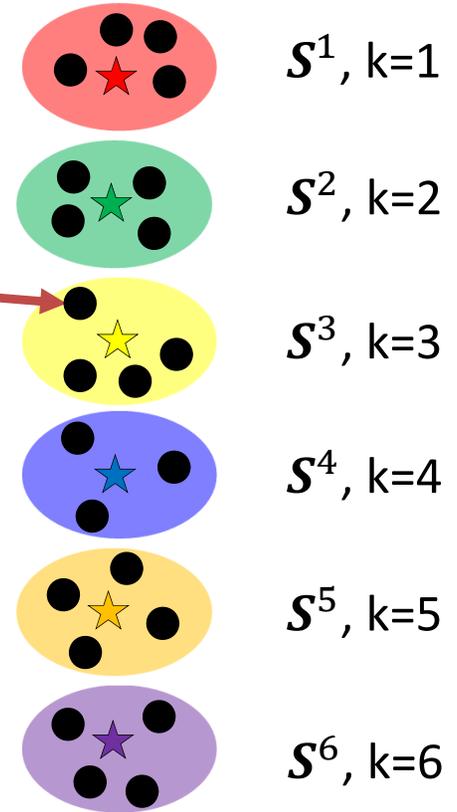
Concept of K-means Clustering

How does K-means clustering work ?



N dimensional
 $A^m = [A_1^m, A_2^m \dots A_N^m]$

Data point A^m



K clusters

Average points W^k

A^m : Data point in cluster S^k

W^k : Average point in cluster S^k

$\| \ \|$: Euclidean distance

k : index of clusters
 m : index of data points

- Cluster:
 - Points with most similar values
 - Has one average point: mean average of nearby data points

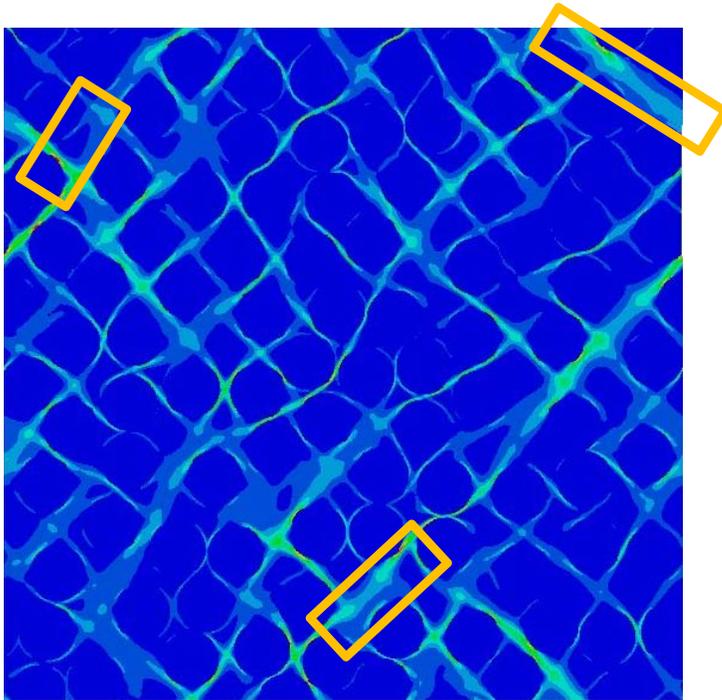
- Objective:
 - Minimize total distance between each average point and the data points within its cluster.

Mathematically: *minimize*:
$$\sum_{k=1}^K \sum_{A^m \in S^k} \|A^m - W^k\|^2$$

K-means clustering for Unidirectional (UD) composite

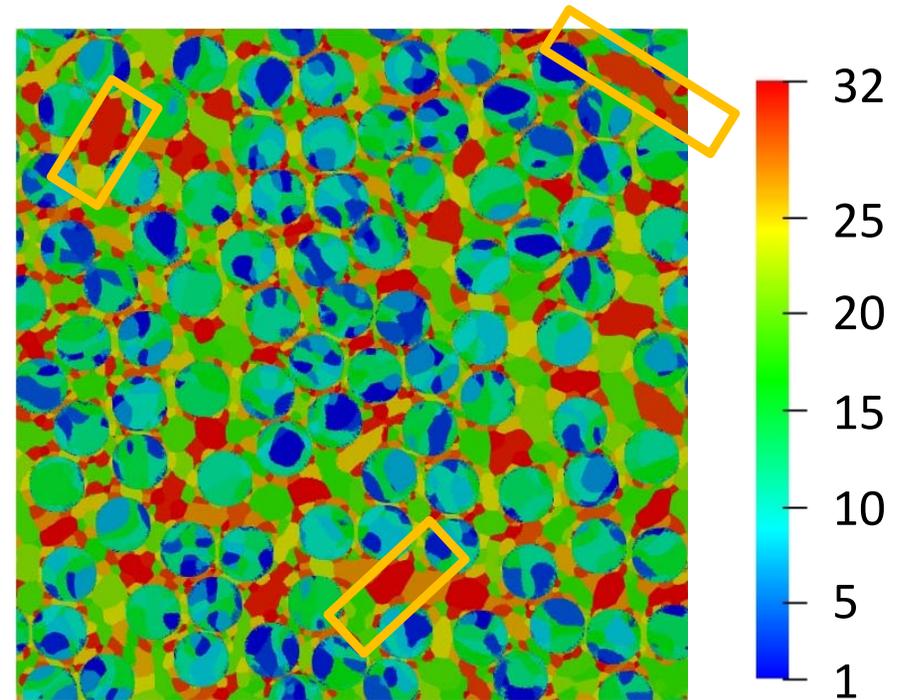
Grouping local material points in the microstructure based on strain responses (or other quantities, such as effective plastic strain)

Strain distribution



2D microstructure with
600 by 600 voxels

Cluster distribution based on strain intensity

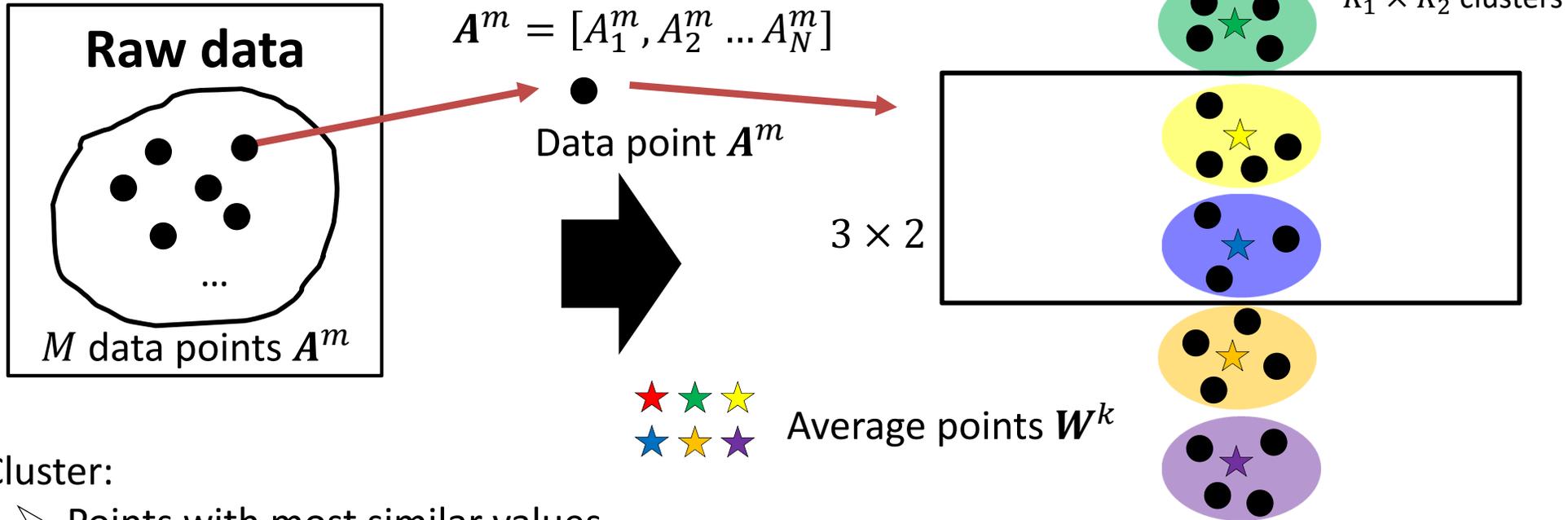


2D microstructure with 32
clusters

- ❖ The strain field, originally represented by 360,000 voxels, is now represented by 32 clusters
- ❖ The strain patterns are adequately captured by the clusters

Self-Organizing Map (SOM) - Concept

How does Self-Organizing Map work? [12,13] N dimensional



- Cluster:
 - Points with most similar values
 - Has one average point: weighted average of nearby data points
- Objective:

Distribute all data points into a **map of $K_1 \times K_2$ clusters** so that the dissimilarity within a cluster is minimized, and the **dissimilarity between clusters with nearby indexes is minimized**

$$\text{minimize: } \sum_{k=[1,1]}^{[K_1, K_2]} \sum_{A^m \in S^k} \sum_{k'=[1,1]}^{[K_1, K_2]} h(\|k - k'\|) \|A^m - W^k\|^2$$

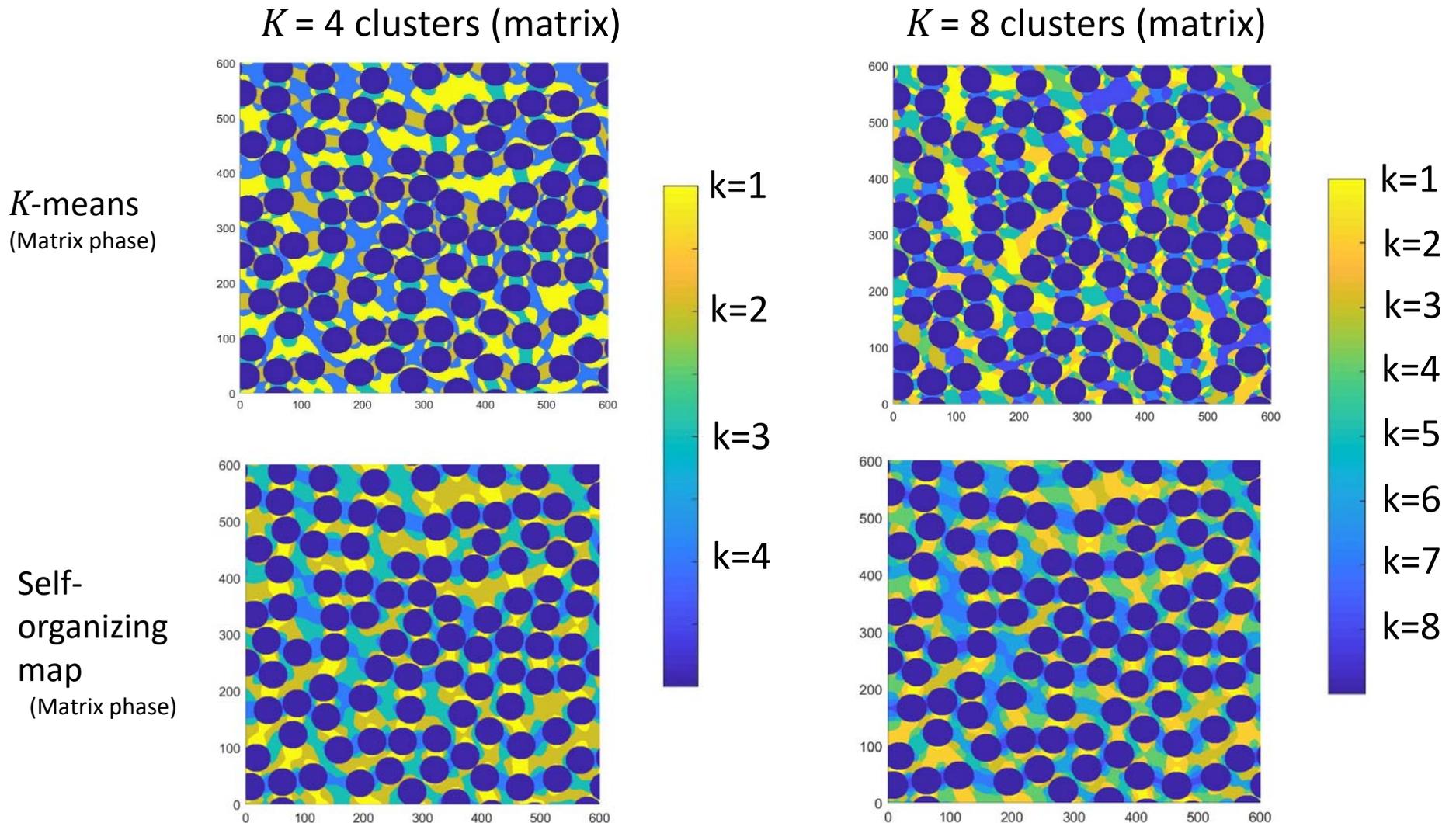
$\|k - k'\|$: Euclidean distance between clusters' indexes

$h(\|k - k'\|)$: Gaussian kernel function



Sample clustering results from K-means and SOM

Unidirectional UD fiber composite (plane strain condition)



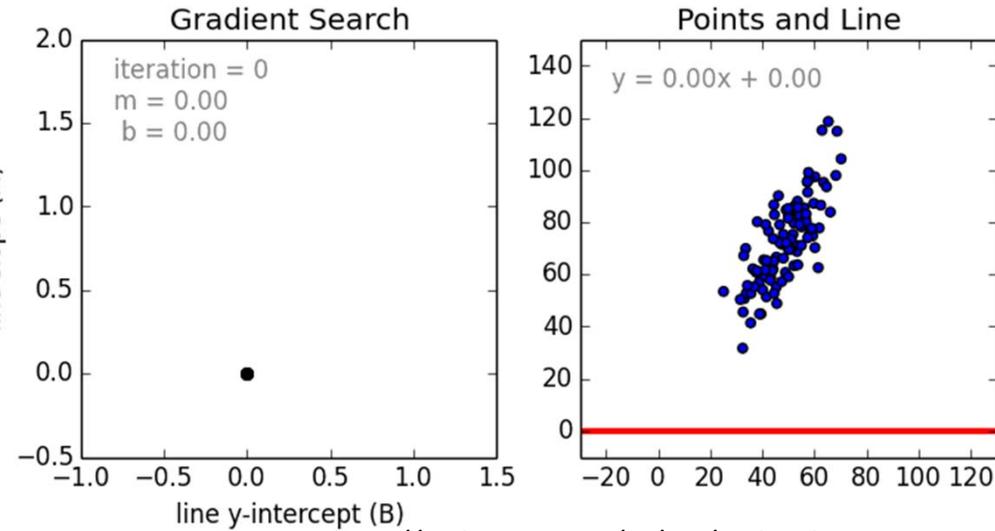
SOM provides orderly ranking of clusters, feature indicators, and physical insights, e.g. strain distribution, damage

Supervised learning

- ❖ Supervised learning establishes the hidden relationship between the input and output data.

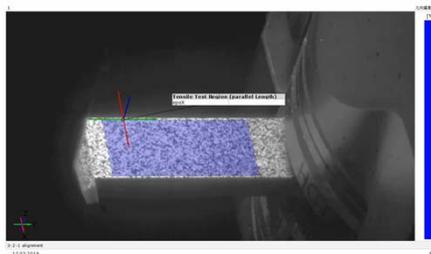


Quick overview of Supervised Learning



Courtesy: <http://aiobserve.com/AI/ML/31.html>

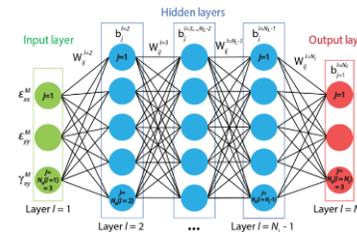
- ❖ Can predict the material law from input and output strain-stress data.



Sample [1]

	ϵ	σ (Mpa)
Data point 1	0.1	20
Data point 2	0.2	38.6
...		

experiment



Training

$$\sigma = NN(\epsilon)$$

Prediction

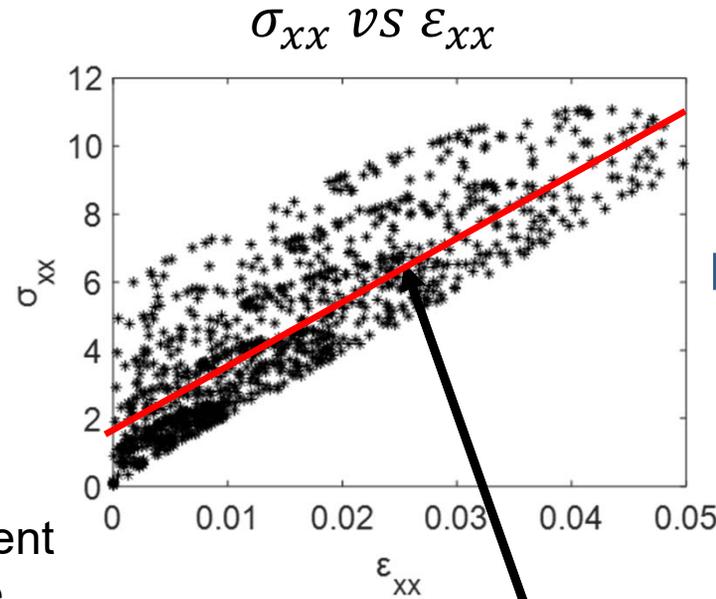
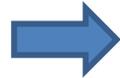
- ❖ Used for – Regression and Classification

Regression

❖ Regression: prediction of response to an input based on *a priori* knowledge of the relationship between input and output data.

Input Data^{1,2}

ϵ	σ
..	..
..	..



Find a relationship between stress and strain
e.g. $\sigma = 2 * \epsilon + 1$

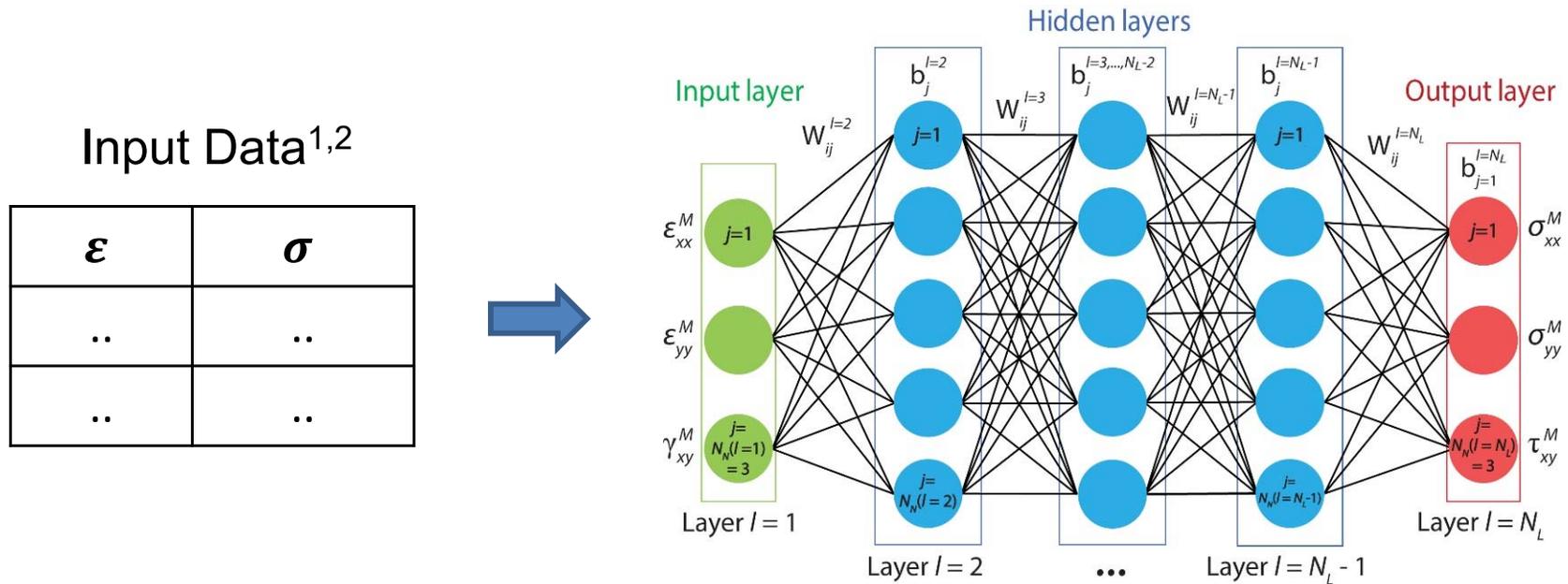
1. Gather data from experiment or high-fidelity microstructure numerical simulation
2. Assume plain strain. ϵ and σ each contain three components

$$\sigma = f(\epsilon) \text{ (hypothesis)}$$

Hypothesis: linear or nonlinear relationship

Regression using Feed Forward Neural Network (FFNN)

- ❖ The relationship between input and output can be explored using FFNN (an ANN with multiple hidden layers)



FFNN training: solving the following optimization problem

$$\text{find} : W_{ij}^{l=2}, b_j^{l=2}, W_{jk}^{l=3}, b_k^{l=3}$$

$$\text{min loss function} : MSE = \frac{1}{N_T \times N_N(l=3)} \sum_{s=1}^{N_T} \sum_{k=1}^{N_N(l=3)} (\sigma_k^{l=3,s} - \sigma_k^{*,l=3,s})^2$$

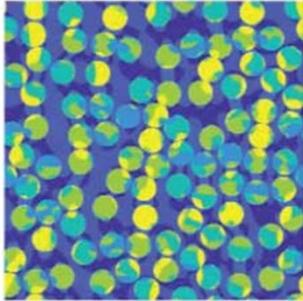
$$\text{where} : \sigma_k^{l=3,s} = \sum_{j=1}^{N_N(l=2)} W_{jk}^{l=3} \left(\mathcal{A} \left(\sum_{i=1}^{N_N(l=1)} W_{ij}^{l=2} \epsilon_i^{M,s} + b_j^{l=2} \right) + b_k^{l=3} \right)$$



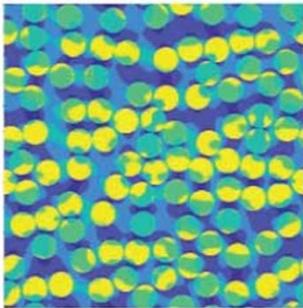
Regression using Convolutional Neural Network (CNN) for inverse modeling

For a microstructure with given **micro-stress** distribution, can CNN predict the **macro-strain**?

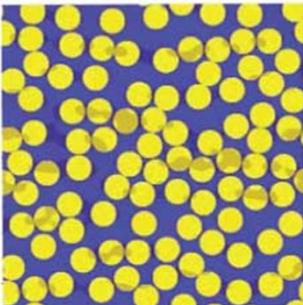
Micro-stress σ_{xx} distribution



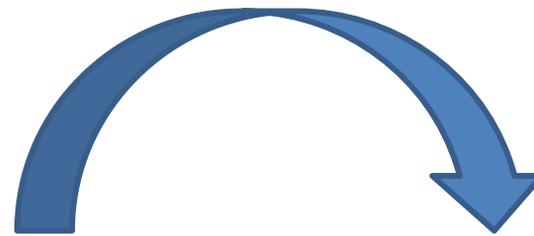
Micro-stress σ_{yy} distribution



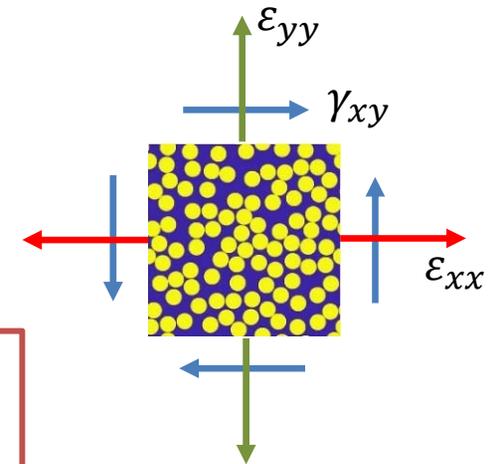
Micro-stress σ_{xy} distribution



Inverse modeling



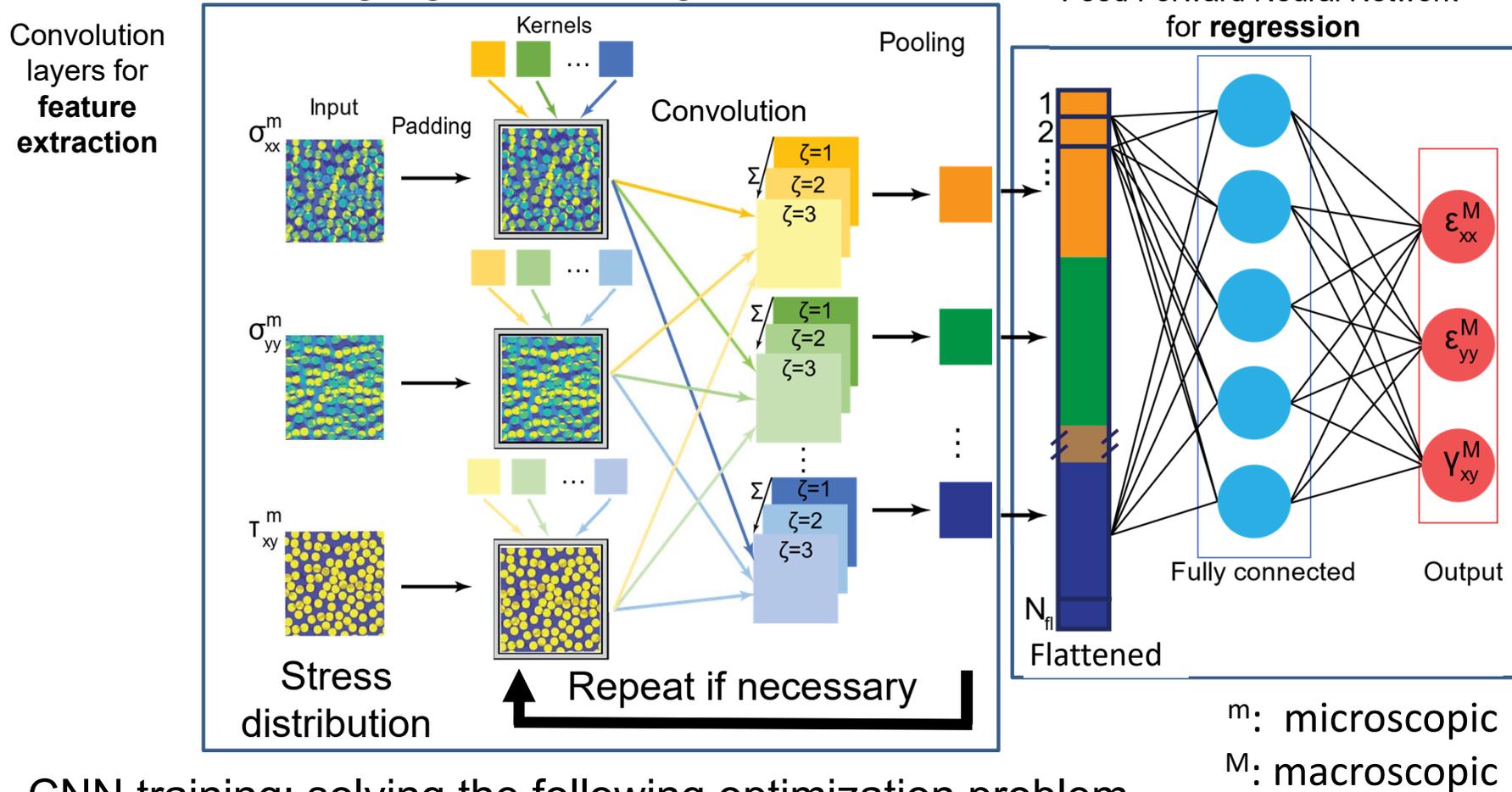
What is the macro-strain ϵ ?



Unidirectional microstructure

Regression using CNN for inverse modeling

Inverse modeling approach to obtain **macroscopic strain** from **microscopic stress distribution** using regression through CNN



CNN training: solving the following optimization problem

$$\text{find : } W_{mn}^l, b_n^l \quad (l = 2, 3 \dots N_L), \text{ in FFNN}$$

$$W_{\epsilon, \sigma}^{\zeta, \kappa, \eta}, b^{\kappa, \eta} \quad (\kappa = 1, 2 \dots N_{kernel}), (\eta = 1, 2 \dots N_{conv}), \text{ in Convolution layers}$$

$$\text{min loss function : } MSE = \frac{1}{N_T} \sum_{s=1}^{N_T} (\epsilon^{M,s} - \epsilon^{*M,s})^2$$

$$\text{where : } \epsilon^{M,s} = \mathcal{F}_{FFNN} \left(\mathcal{F}_{flatten} \left(\mathcal{F}_{pcp}^{N_{conv}} \left(\dots \mathcal{F}_{pcp}^2 \left(\mathcal{F}_{pcp}^1 \left(\sigma^s(\alpha, \beta) \right) \right) \dots \right) \right) \right)$$

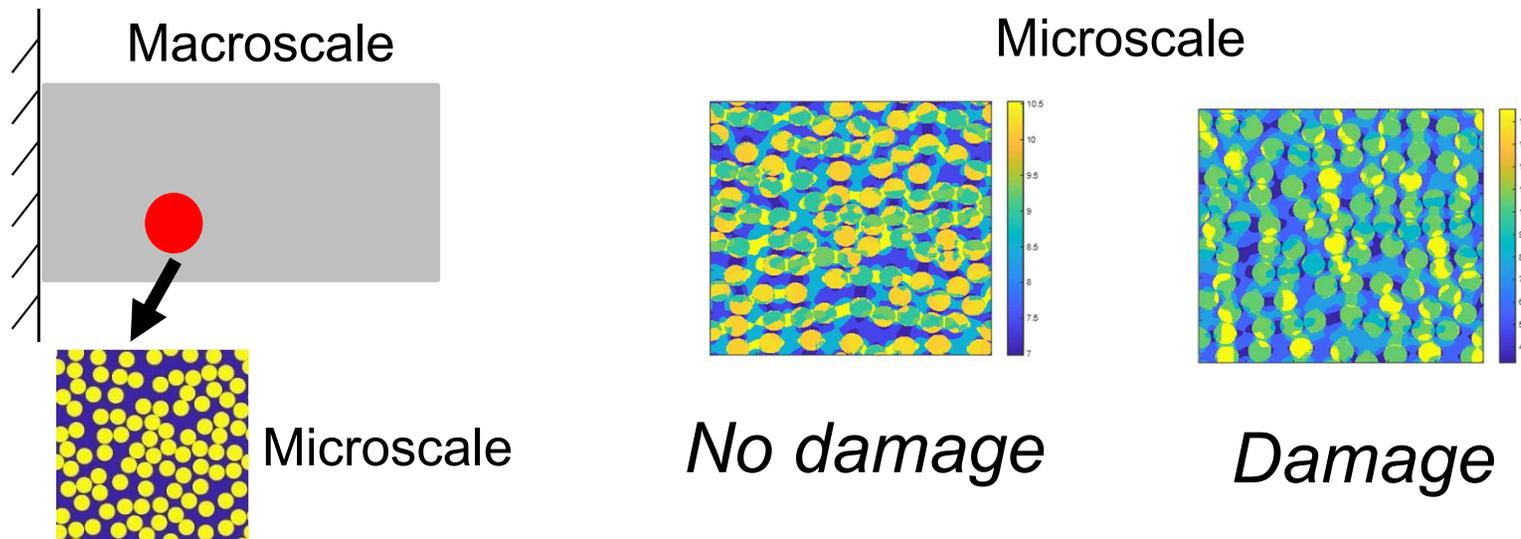
Classification of damage

❖ Classification is a process of predicting the known class of given data points.

E.g., classify the state of the microstructure as “no damage” or “damage” based on local stress distribution

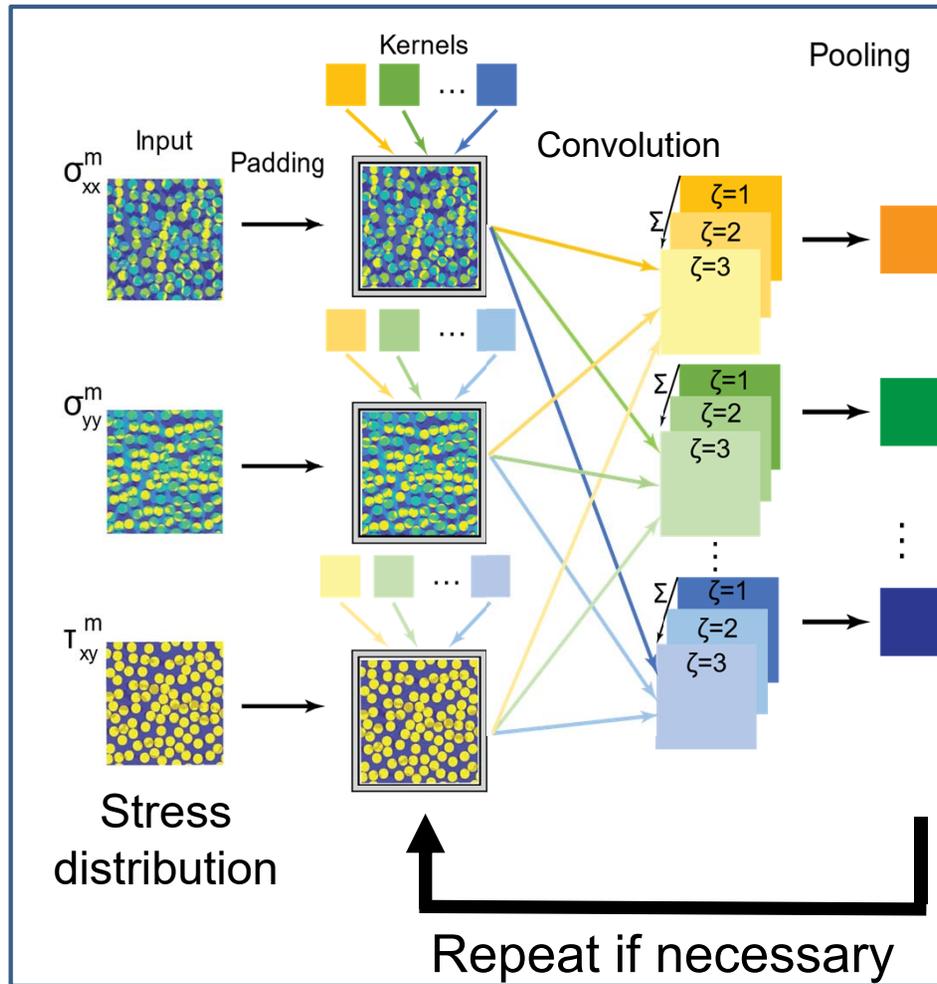
- Microscale material point damage is defined as: for any material point in the **microscale** domain, if the **micro-stress** exceeds certain threshold, the **micro** material point is damaged
- **Macroscale material point** damage is defined as:
 - 1) $P_D > P_{ND}$, damage in the microstructure
 - 2) $P_D < P_{ND}$, no damage in the microstructure

*P is probability

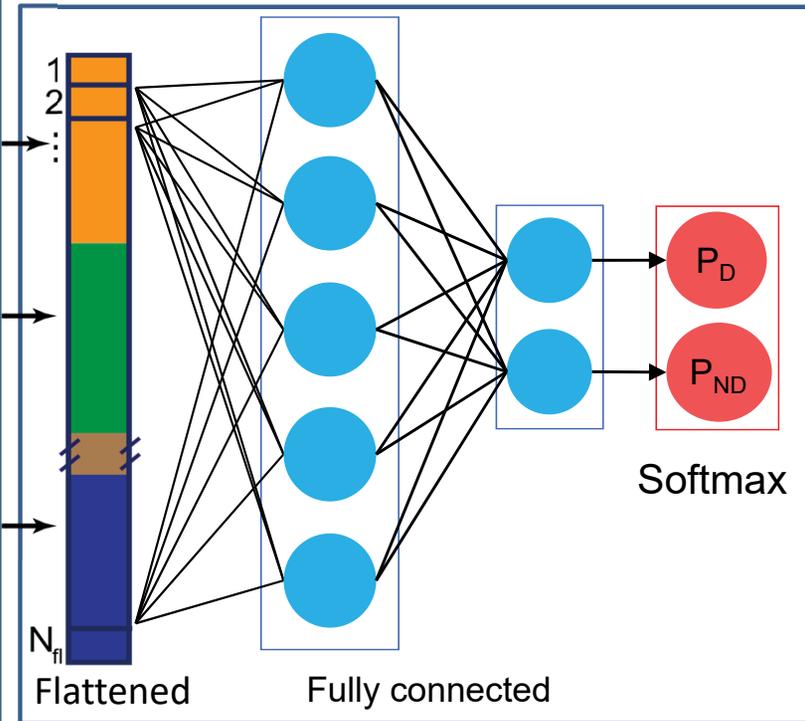


Application of CNN for damage classification

Convolution layers for feature extraction



Feedforward Neural Network with Softmax layer



P_D : probability of damage
 P_{ND} : probability of no damage

Macroscale material point damage is defined:

- 1) $P_D > P_{ND}$, damage in the microstructure
- 2) $P_D < P_{ND}$, no damage in the microstructure

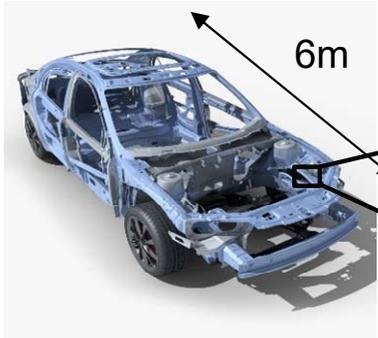


Outline

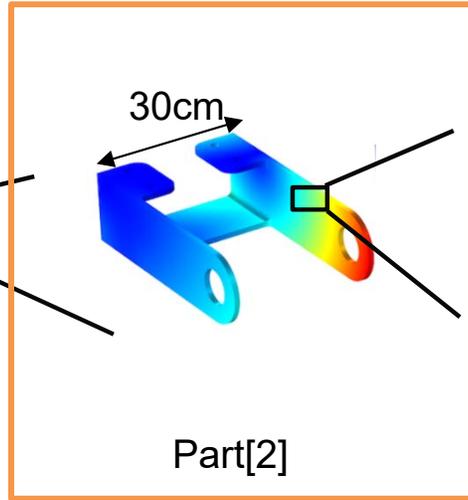
1. Motivation: source of data in mechanical science and engineering
2. Mechanistic Machine Learning (MML) for mechanical science and engineering
 - Interpretation of the data
 - Relevant concepts in data science
 - Introduction to different Machine Learning (ML) methods
 - a. Unsupervised learning
 - b. Supervised learning
- 3. Applications of ML methods**
 1. Topology optimization
 1. Feed Forward Neural Network (FFNN)
 2. FFNN+ Convolutional Neural Network (CNN)
 2. Adolescent Idiopathic Scoliosis
 1. FFNN
 2. Physics Guided Neural Network (PGNN)
4. Why we need reduced order models/methods (ROM)
5. Summary and conclusions
6. References



Multiple length scales composite systems design & Optimization

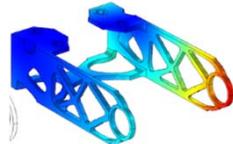


Car frame[1]

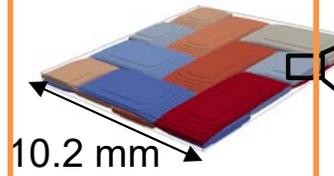


Part[2]

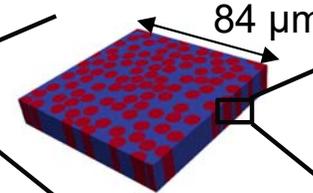
Topology



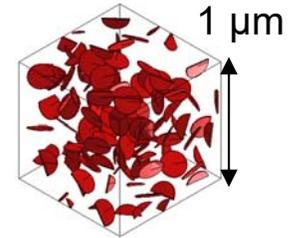
Density distribution: $\rho(x)$



Woven composite



UD fiber composite



Nanoscale reinforced polymer matrix

Multiscale Design:

Woven thickness: h_{layer} Fiber radius: r_{fiber} MoS₂ volume fraction: V_{MoS2}
Layer angle: α_{layer} Fiber volume fraction: V_{fiber}

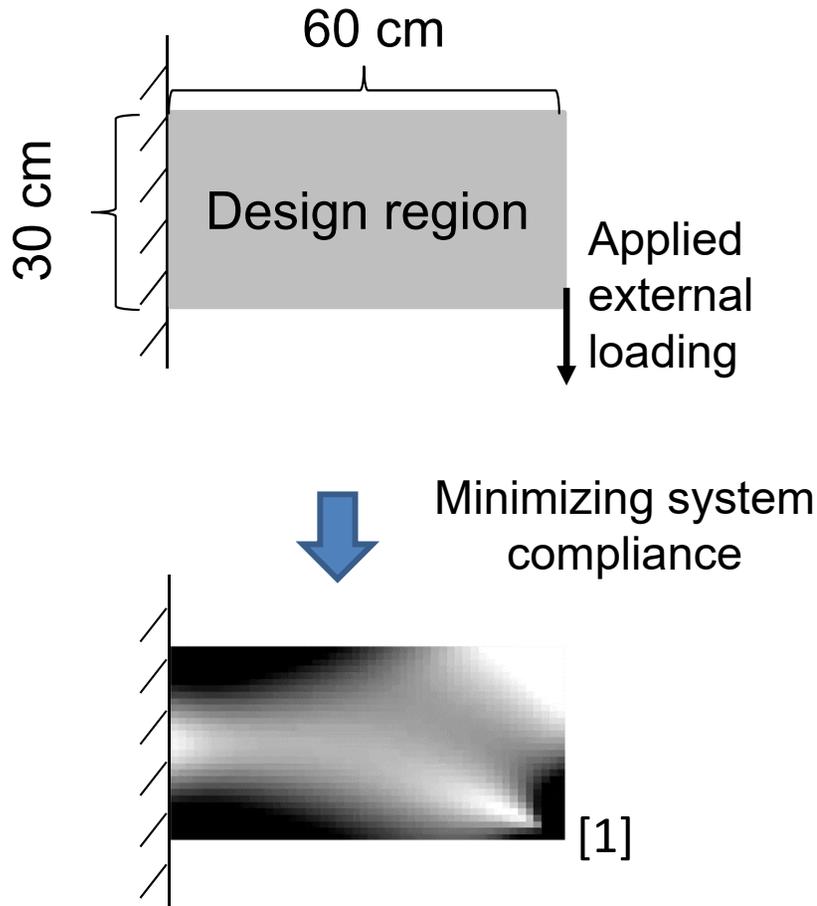
Name	Part	Woven composite	UD composite	MoS2 polymer	Total
Length scale	cm	mm	μm	μm	-
Number of elements	10,000	40,000	360,000	90,000	1.296×10 ¹⁹
Approximate optimization iterations	200	200	200	100	-
Total calculation cost	-	-	-	-	A tremendous number

[1]<https://www.cgtrader.com/3d-models/vehicle/part/car-frame-03>

[2]<https://www.comsol.com/blogs/performing-topology-optimization-with-the-density-method/>

Topology optimization (TopOpt)

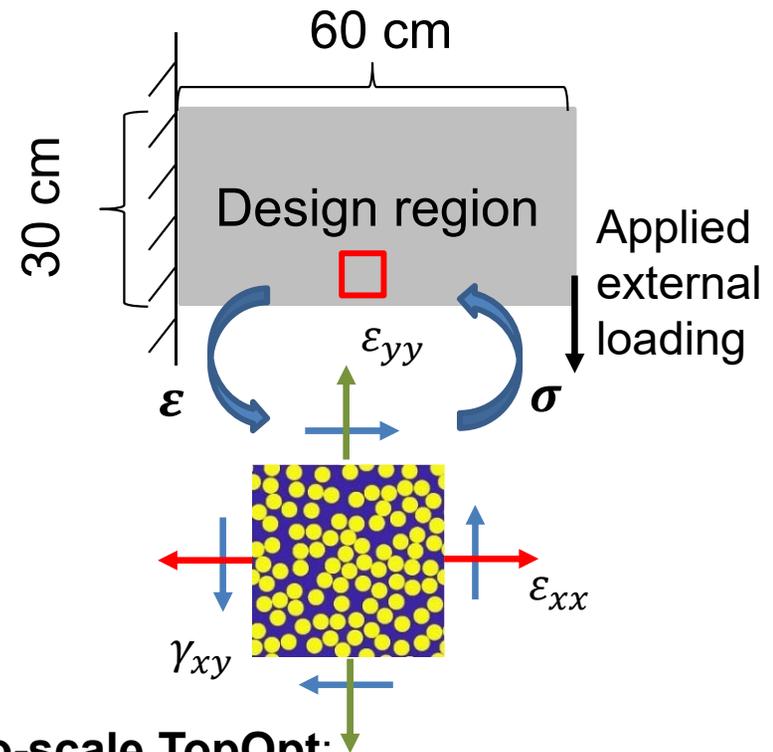
Single-scale topology optimization



- Homogenous material assumed
- No microstructure
- Only elastic responses considered

[1] Sigmund, O. (2001). A 99 line topology optimization code written in Matlab. *Structural and multidisciplinary optimization*, 21(2), 120-127.

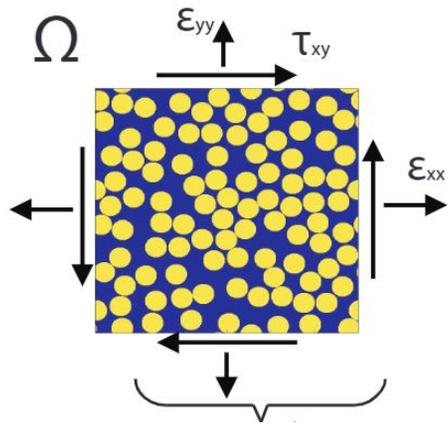
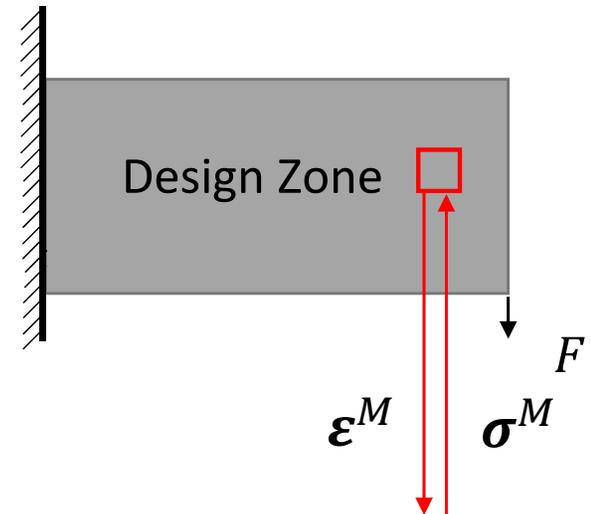
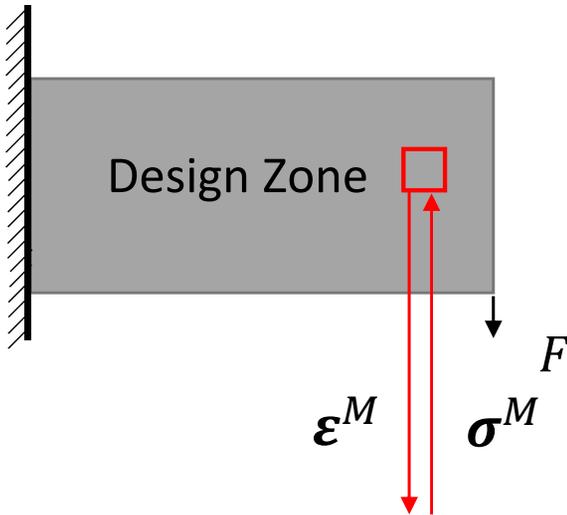
Microstructure-based topology optimization is a two-scale problem



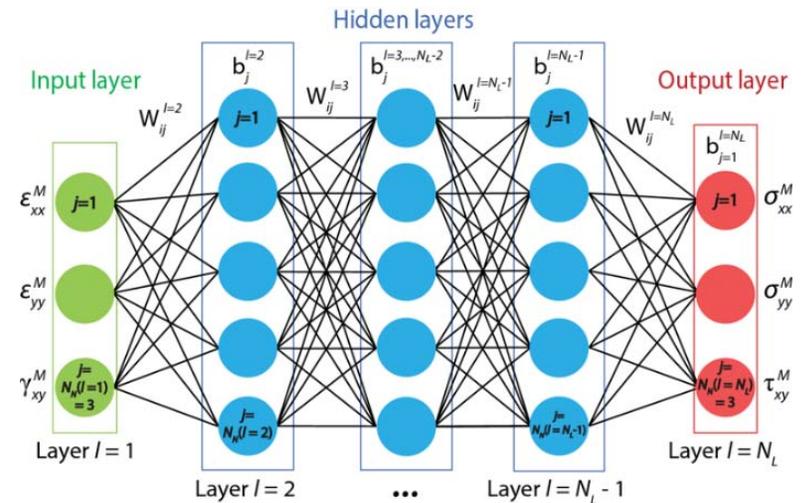
Two-scale TopOpt:

- Microstructures in all material points
- Design of microstructures and structure topology
- Evaluation of microstructure is time consuming during design iterations
- Can FFNN and CNN approximate microstructure responses efficiently and accurately?

Topology optimization with FFNN



Representative Volume Element (RVE)



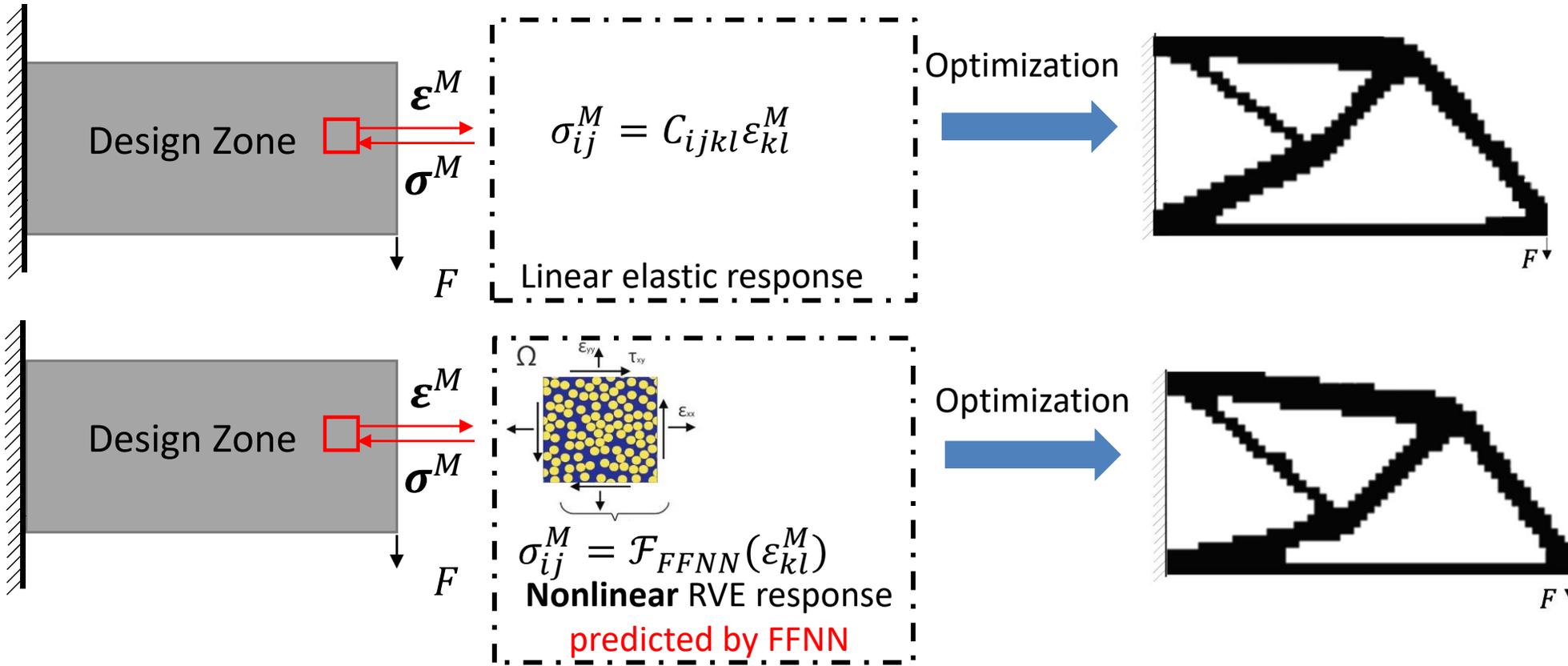
Feedforward Neural Network

FFNN approximates microstructure responses almost instantaneously

To be presented by Hengyang Li, 7/29/2019, 4:50-5:10pm, Room 202

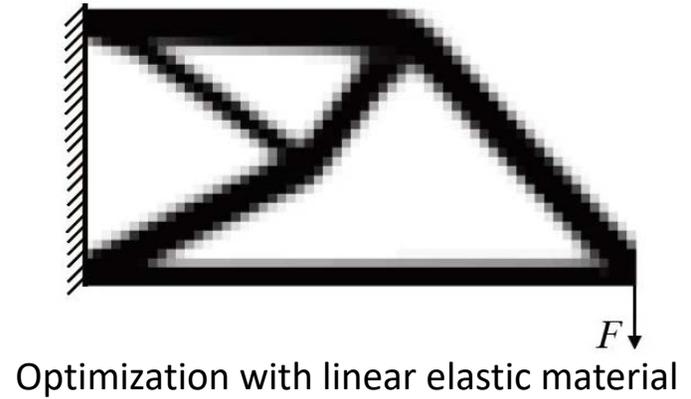
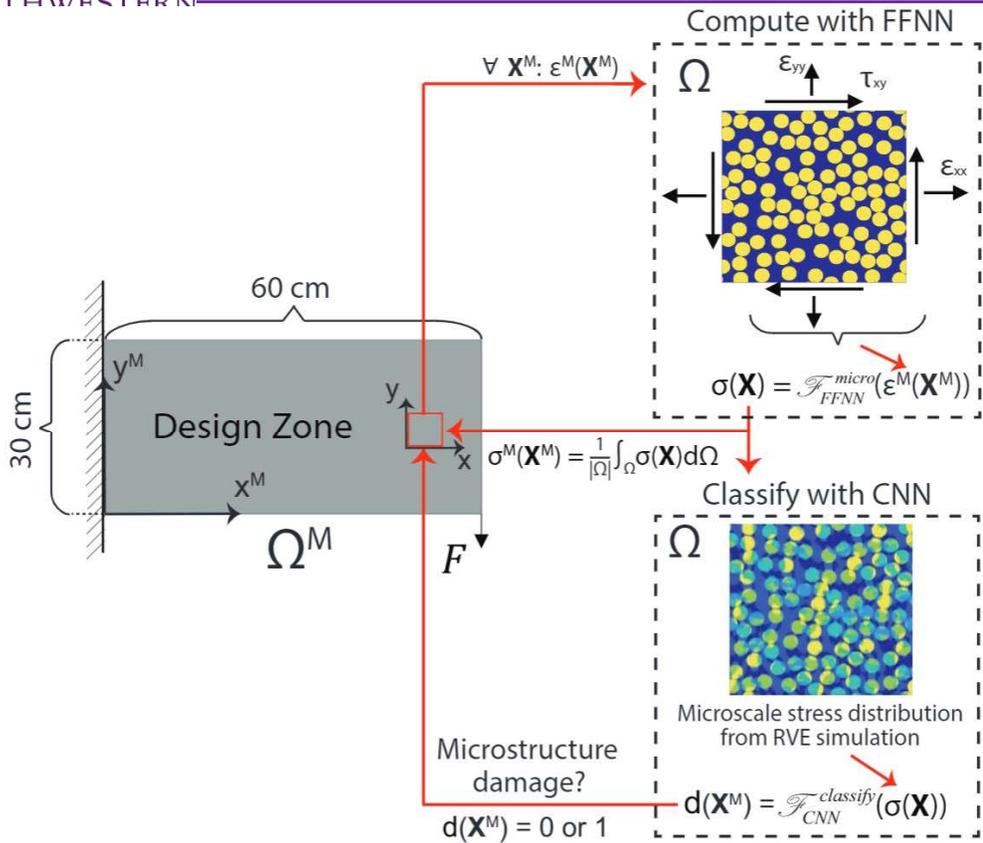


TopOpt with FFNN for nonlinear elastic materials



	Linear material	Nonlinear FEM-FEM two scale	Nonlinear FEM-FFNN two scale
Initial compliance $N \cdot cm$	295.0	-	375.0
Optimized compliance $N \cdot cm$	28.0	-	38.0
Optimization calculation time s	338	220×10^6	472
Factor of speed-up over FE-FE	-	-	280,255

FEM: finite element method



Linear material

Nonlinear FEM-FEM two scale

Nonlinear FEM-(FFNN+CNN) two scale

	Linear material	Nonlinear FEM-FEM two scale	Nonlinear FEM-(FFNN+CNN) two scale
Initial compliance $N \cdot cm$	295.0	-	295
Optimized compliance $N \cdot cm$	30	-	31
Optimization calculation time s	12.6	660×10^6	14.5
Factor of speed-up over FE-FE	-	-	45×10^6

FEM: finite element method



Outline

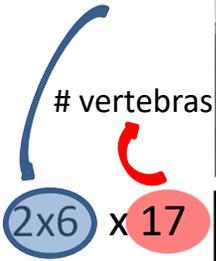
1. Motivation: source of data in mechanical science and engineering
2. Mechanistic Machine Learning (MML) for mechanical science and engineering
 - Interpretation of the data
 - Relevant concepts in data science
 - Introduction to different Machine Learning (ML) methods
 - a. Unsupervised learning
 - b. Supervised learning
- 3. Applications of ML methods**
 1. Topology optimization
 1. Feed Forward Neural Network (FFNN)
 2. FFNN+ Convolutional Neural Network (CNN)
 - 2. Adolescent Idiopathic Scoliosis**
 1. FFNN
 2. Physics Guided Neural Network (PGNN)
4. Why we need reduced order models/methods (ROM)
5. Summary and conclusions
6. References



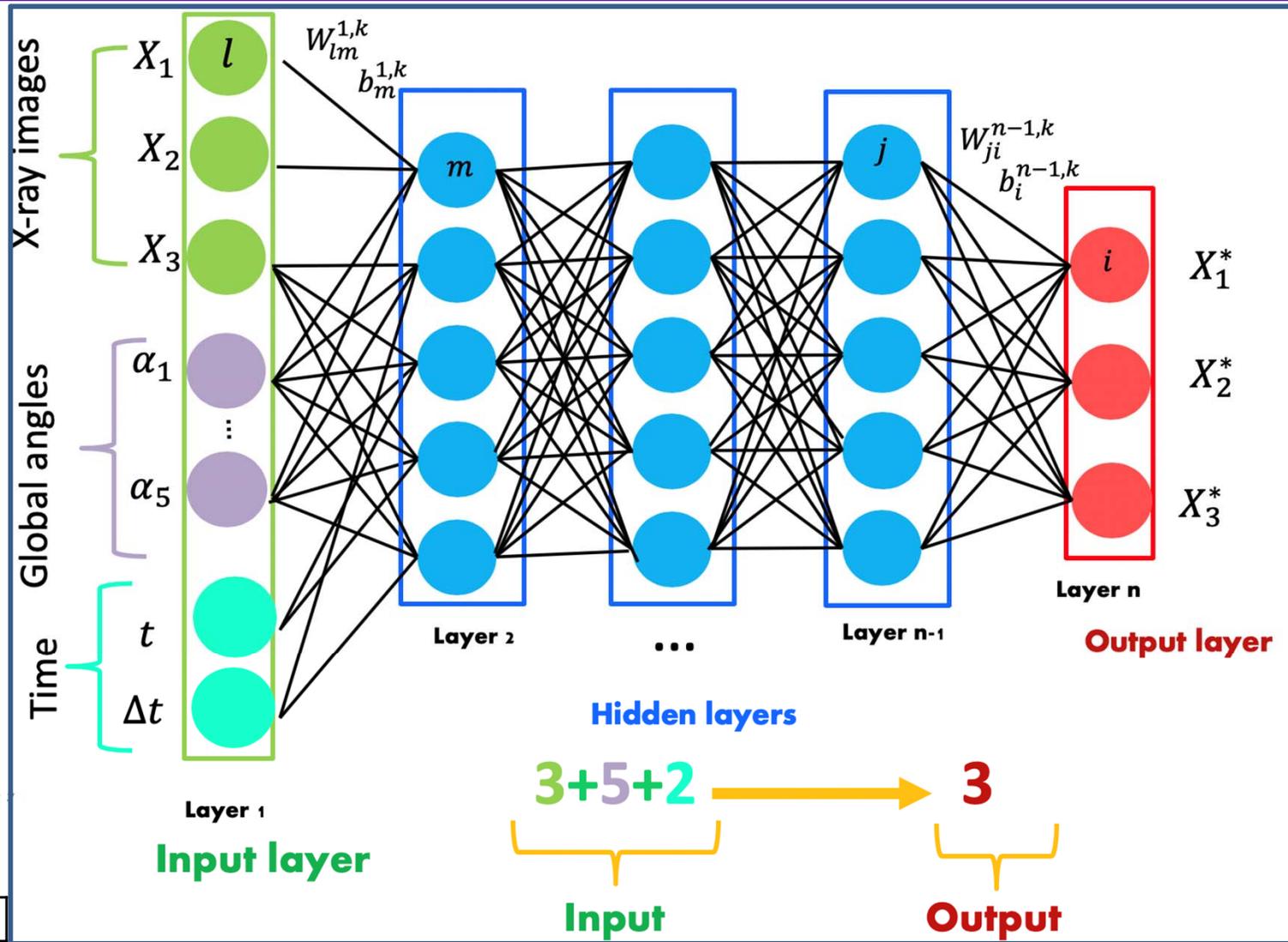
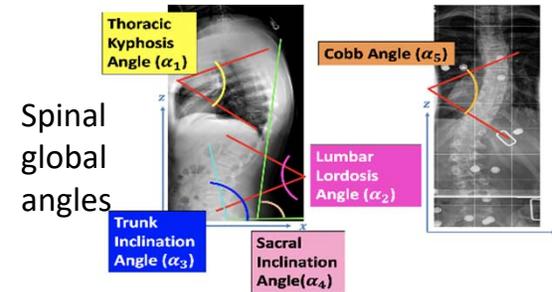
Data-driven approach in predicting Adolescent Idiopathic Scoliosis (AIS)

NORTHWESTERN UNIVERSITY

6 Landmarks



6 Landmarks



Features	Data points (l)					
	1	2	3	.	.	Ns
X						
α
t
Δt						
X*						

X = Vector of input coordinates of a landmark $[X_1 \ X_2 \ X_3]$
 α = Global angle vector $[\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4 \ \alpha_5]$
 t = Age of the patient.
 Δt = age variance between target age and current age (month).
 X^* = Vector of output co-ordinates of a landmark $[X_1^* \ X_2^* \ X_3^*]$.
Ns = Total number of landmarks = $2 \times 6 \times 17 = 204$

Using all features to train the Neural Network for all $l=1, \dots, Ns$

To be presented by Mahsa Tajdari, 7/31/2019, 2:40-3:00pm, Room 208

Assume linear relationship between effective stress and **growth rate** (\dot{X}_{mn}) between time $m\Delta t$ and $n\Delta t$

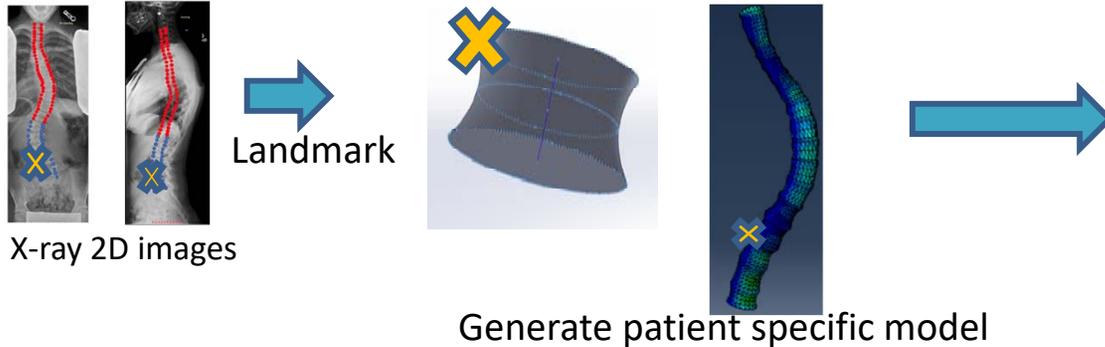
$$\dot{X}_{mn} = (m - n)[A(1 + B \times \sigma_{eff}^n)]$$

Δt : unit of time (month)

σ_{eff}^n : effective stress at time $n\Delta t$

A ($month^{-1}$) and **B** (MPa^{-1}) are patient specific constants that are calibrated inside the NN

Calculating Hyper Parameters in NN



$$MSE_{X^{*,1}} = MSE_1^1 + MSE_2^1 + MSE_3^1$$

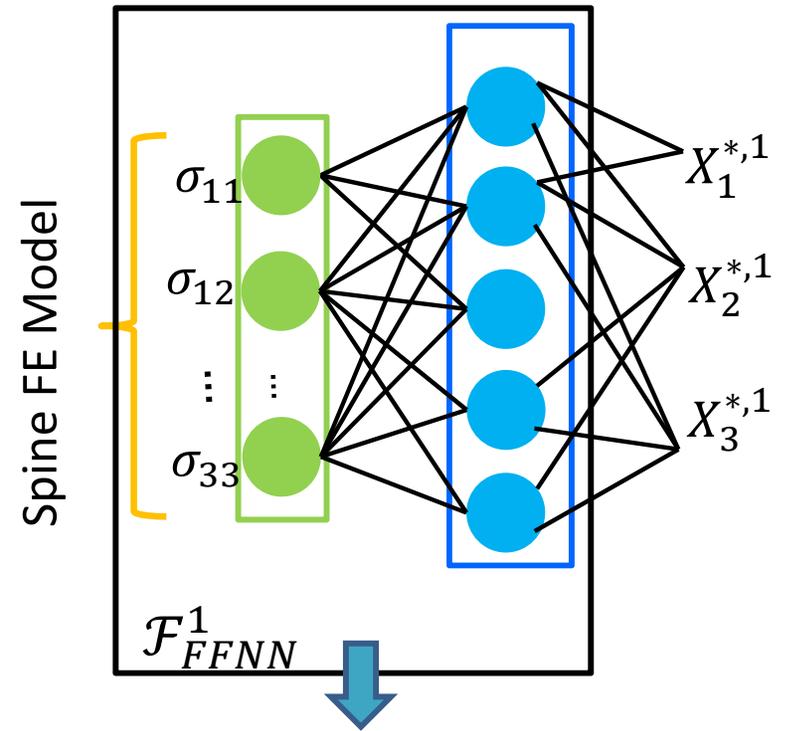
$$MSE_1^1 = MSE(X^{*,1} - \tilde{X}^{*,1})$$

$$MSE_2^1 = MSE(\dot{X}_{mn} - ((m - n)[A_{TH}(1 + B_{TH} \times \sigma_{eff}^n(TH))]))$$

$$MSE_3^1 = MSE(\dot{X}_{mn} - ((m - n)[A_{LU}(1 + B_{LU} \times \sigma_{eff}^n(LU))]))$$

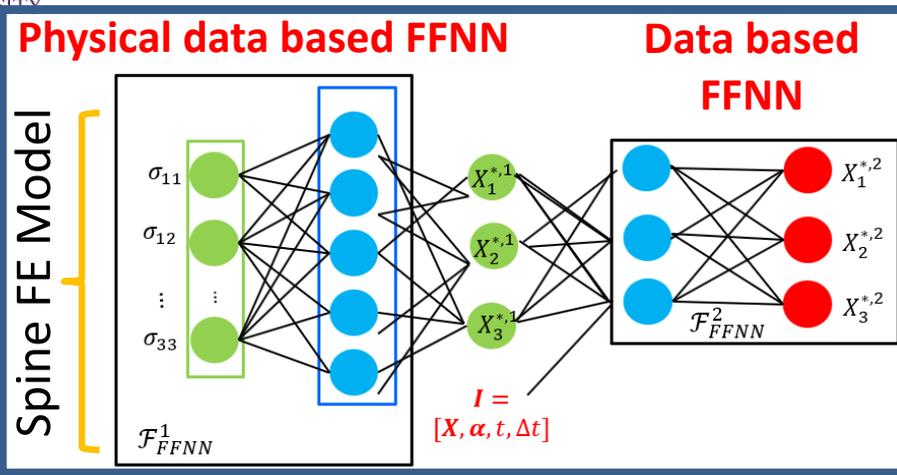
A_{TH} ($month^{-1}$) and **B_{TH}** (MPa^{-1}) patient specific constants for Thoracic vertebrae

A_{LU} ($month^{-1}$) and **B_{LU}** (MPa^{-1}) patient specific constants for Lumbar vertebrae



Hyper parameters (**A_{TH}**, **B_{TH}**, **A_{LU}**, **B_{LU}**) will be calculate inside NN

Composite NN using Multi-fidelity data



Physical Equation

$$\left((m - n) \left[A_{TH} \left(1 + B_{TH} \times \sigma_{eff}^n (TH) \right) \right] \right) - \dot{X}_{mn} = 0$$

$$\left((m - n) \left[A_{LU} \left(1 + B_{LU} \times \sigma_{eff}^n (LU) \right) \right] \right) - \dot{X}_{mn} = 0$$

α = Global angle vector [$\alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5$]
 t = Age of the patient.

Δt = age variance between target age and current age (month).

X^* = Vector of output co-ordinates of a landmark [$X_1^* X_2^* X_3^*$]

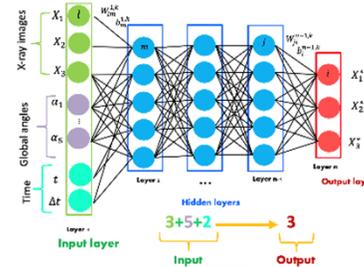
X^m : position of data point in m^{th} month
 X^n : position of data point in n^{th} month

$$\dot{X}_{mn} = \frac{\sqrt{(X_1^m - X_1^n)^2 + (X_2^m - X_2^n)^2 + (X_3^m - X_3^n)^2}}{\left(\sqrt{(X_1^n)^2 + (X_2^n)^2 + (X_3^n)^2} \right)}$$

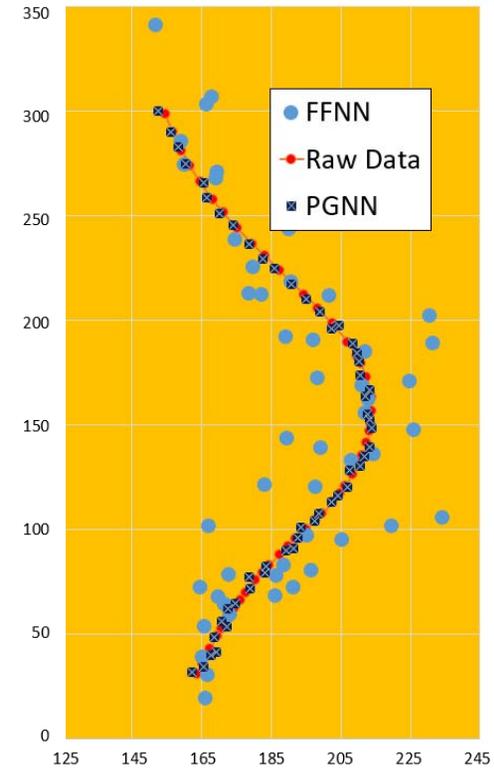
Raw Data



FFNN



PGNN



Relative Error
 Data based FFNN: 18.5%
PGNN: 4.63%



Outline

1. Motivation: source of data in mechanical science and engineering
2. Mechanistic Machine Learning (MML) for mechanical science and engineering
 - Interpretation of the data
 - Relevant concepts in data science
 - Introduction to different Machine Learning (ML) methods
 - a. Unsupervised learning
 - b. Supervised learning
3. Applications of ML methods
 1. Topology optimization
 1. Feed Forward Neural Network (FFNN)
 2. FFNN+ Convolutional Neural Network (CNN)
 2. Adolescent Idiopathic Scoliosis
 1. FFNN
 2. Physics Guided Neural Network (PGNN)
4. **Why we need reduced order models/methods (ROM)**
5. Summary and conclusions
6. References



Rich database of mechanical response information are necessary for training various Neural Networks

- Multiscale design and optimization is not feasible with direct microstructure responses calculation with Finite Element Method (FEM)
- Well-trained **NNs accelerates** microstructure and structure **design process**, e.g. Topology Optimization
- **Material microstructure responses database** is required for the training process.
- The database includes:
 - Macro-strain and macro-stress pairs
 - Micro-stress distribution and macro-strain pairs
 - Other microstructure quantities of interest

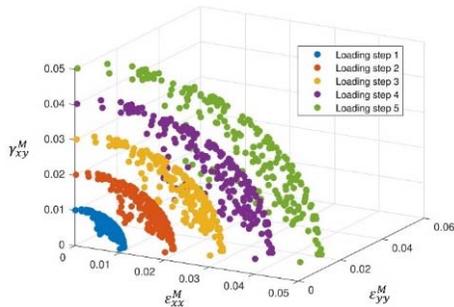
The gap:

- Microstructure response simulation can be expensive using FEM
- Rich database requires a lot of runs of microstructure simulation

Cost of microstructure responses database generation

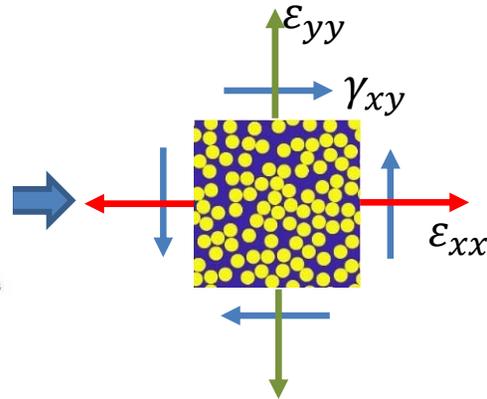
1,000 load cases for training a 2D hyper elastic problem:

External loading states

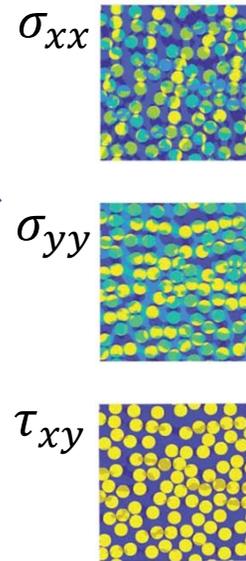


$$\boldsymbol{\epsilon}^n = [\epsilon_{xx}^n, \epsilon_{yy}^n, \gamma_{xy}^n]$$

$$n = 1, 2, \dots, 1000$$

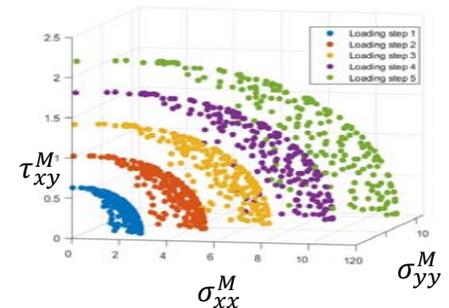


Microscopic stress



600 x 600 x 3 x 1000

Averaged stress



$$\boldsymbol{\sigma}^n = [\sigma_{xx}^n, \sigma_{yy}^n, \tau_{xy}^n]$$

$$n = 1, 2, \dots, 1000$$

Running 1,000 microstructure simulation is expensive:

Microstructure simulation method	Total simulation time (s)
FFT	3.01×10^5
FEM	2.04×10^7

HPC is needed

Approximate material responses using FFNN

External loading states

Data point	Feature
1	ϵ^1
2	ϵ^2
...	
n	ϵ^n

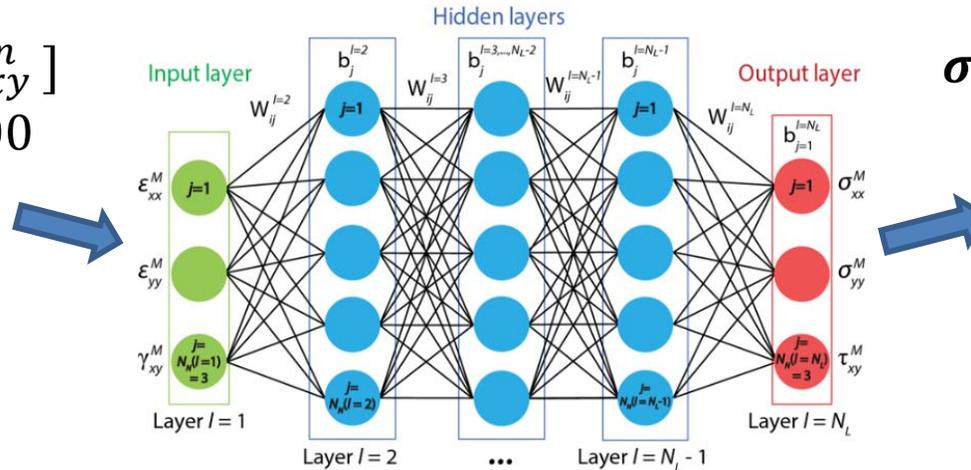
Averaged stress

Data point	Feature
1	σ^1
2	σ^2
...	
n	σ^n

FFNN with well-trained parameters

$$\epsilon^n = [\epsilon_{xx}^n, \epsilon_{yy}^n, \gamma_{xy}^n]$$

$$n = 1, 2, \dots, 1000$$



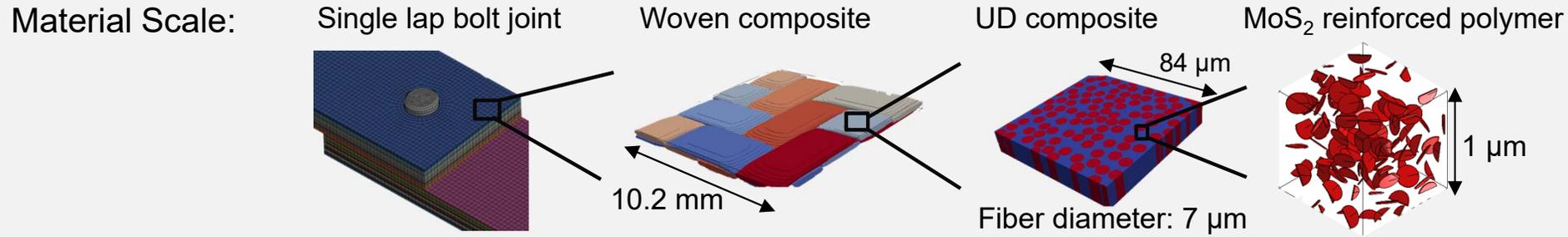
$$\sigma^n = [\sigma_{xx}^n, \sigma_{yy}^n, \tau_{xy}^n]$$

$$n = 1, 2, \dots, 1000$$

How to generate microstructure responses efficiently?



Curse of dimensionality in complex material systems



Typical number of elements:	1×10^6	2×10^6	36×10^6	64×10^3
Clusters used:	10 clusters	96 clusters	10 clusters	10 clusters

FEM

$$1.38 \times 10^{25} \begin{bmatrix} \{\mathbf{K}\}^{(1)(1)} & \{\mathbf{K}\}^{(1)(2)} & \dots & \{\mathbf{K}\}^{(1)(4)} \\ \{\mathbf{K}\}^{(2)(1)} & \{\mathbf{K}\}^{(2)(2)} & \dots & \{\mathbf{K}\}^{(2)(4)} \\ \vdots & \vdots & \ddots & \vdots \\ \{\mathbf{K}\}^{(4)(1)} & \{\mathbf{K}\}^{(4)(2)} & \dots & \{\mathbf{K}\}^{(4)(4)} \end{bmatrix} \begin{bmatrix} \{\delta \mathbf{d}\}^{(1)} \\ \{\delta \mathbf{d}\}^{(2)} \\ \vdots \\ \{\delta \mathbf{d}\}^{(4)} \end{bmatrix} = \begin{bmatrix} \{\mathbf{r}\}^{(1)} \\ \{\mathbf{r}\}^{(2)} \\ \vdots \\ \{\mathbf{r}\}^{(4)} \end{bmatrix}$$

1.38×10^{25}

Strong interaction between scales

DoFs using FEM:
 $3 \left(N_n^{(1)} + N_i^{(1)} N_n^{(2)} + \dots + N_i^{(1)} N_i^{(2)} N_i^{(3)} N_n^{(4)} \right)$

Number of elements using FEM:
 $1 \times 10^6 * 2.6 \times 10^6 * 36 \times 10^6 * 64 \times 10^3 = 6 \times 10^{25}$ elements

N_i : number of integration points
 N_n : number of nodes
 Superscripts indicate scale level

Extremely large problem

MCA

$$639420 \begin{bmatrix} \{\mathbf{M}\}^{(1)(1)} & \{\mathbf{M}\}^{(1)(2)} & \dots & \{\mathbf{M}\}^{(1)(4)} \\ \{\mathbf{M}\}^{(2)(1)} & \{\mathbf{M}\}^{(2)(2)} & \dots & \{\mathbf{M}\}^{(2)(4)} \\ \vdots & \vdots & \ddots & \vdots \\ \{\mathbf{M}\}^{(4)(1)} & \{\mathbf{M}\}^{(4)(2)} & \dots & \{\mathbf{M}\}^{(4)(4)} \end{bmatrix} \begin{bmatrix} \{\delta \boldsymbol{\varepsilon}\}^{(1)} \\ \{\delta \boldsymbol{\varepsilon}\}^{(2)} \\ \vdots \\ \{\delta \boldsymbol{\varepsilon}\}^{(4)} \end{bmatrix} = \begin{bmatrix} \{\mathbf{r}\}^{(1)} \\ \{\mathbf{r}\}^{(2)} \\ \vdots \\ \{\mathbf{r}\}^{(4)} \end{bmatrix}$$

639420

DoFs using MCA:
 $6 \left(N_c^{(1)} + \dots + N_c^{(1)} N_c^{(2)} N_c^{(3)} N_c^{(4)} \right)$

Number of clusters using MCA:
 $10 * 96 * 10 * 10 = 9.6 \times 10^4$ clusters

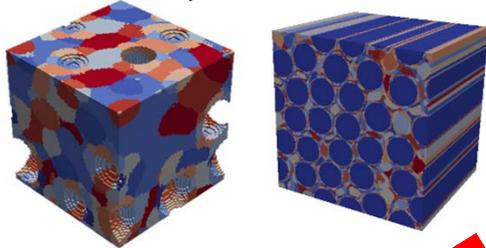
N_c : number of clusters
 $N_c \ll N_n$ at each scale

Solvable on small HPC/single PC

- **Objective:** Efficient and accurate homogenization of **nonlinear history dependent** heterogeneous materials with **complex microstructure**.

Data-driven order reduction

Group points in the MVE that are mechanically similar



Mechanistic prediction

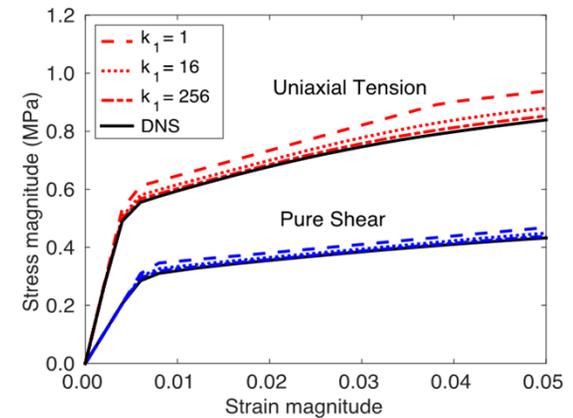
- Lippmann-Schwinger integral equation
- Micromechanics mean-field theory

Self-consistent Clustering Analysis (SCA)

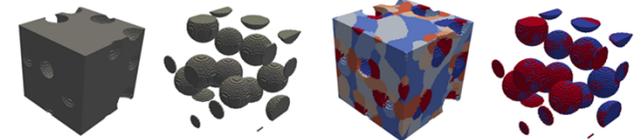
1. Liu, Z., Bessa, M. A., & Liu, W. K. (2016). Computer Methods in Applied Mechanics and Engineering
2. Liu, Z., Fleming, M., & Liu, W. K. (2018). Computer Methods in Applied Mechanics and Engineering
3. Bessa, M. A., Bostanabad, R., Liu, Z., Hu, A., Apley, D. W., Brinson, C., ... & Liu, W. K. (2017). Computer Methods in Applied Mechanics and Engineering
4. Liu, Z., Kafka, O. L., Yu, C., & Liu, W. K. (2018). In Advances in Computational Plasticity
5. Tang, S., Zhang, L., & Liu, W. K. (2018). Computational Mechanics
6. Kafka, O. L., Yu, C., Shakoor, M., Liu, Z., Wagner, G. J., & Liu, W. K. (2018). JOM
7. Shakoor, M., Kafka, O. L., Yu, C., & Liu, W. K. (2018). Computational Mechanics
8. Li, H., Kafka, O. L., Gao, J., Yu, C., Nie, Y., Zhang, L., ... & Tang, S. (2019). Computational Mechanics
9. Zhang, L., Tang, S., Yu, C., Zhu, X., & Liu, W. K. (2019). Computational Mechanics
10. Gao, J., Shakoor, M., Jinnai, H., Kadowaki, H. Seta, E., Liu, W. K. An Inverse Modeling Approach for Predicting Filled Rubber Performance. (2019) Computer Methods in Applied Mechanics and Engineering

Numerical verification

3-Dimension



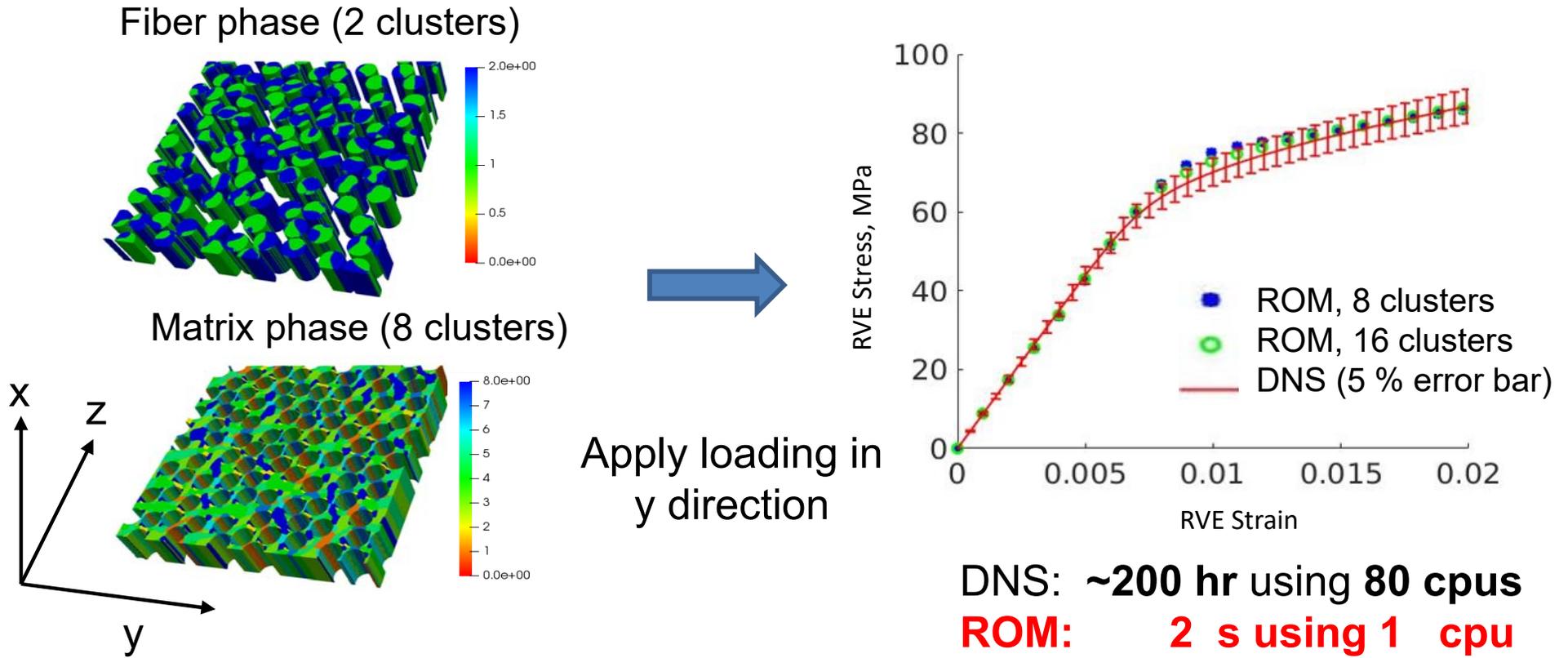
FEM Model Reduced Order Model (ROM)



	System Complexity	Computational time	Speed-up
FEM	80x80x80	25.7 hr (24 cores)	1
ROM	16	2 s	1x10⁶
(SCA)	256	50 s	5x10⁴

To be covered in Lecture 2

Why use reduced order modeling for data generation?



- Material design requires a large database of microstructure response information
- Reduced order modeling (ROM) allows fast data generation for:
 - Different heterogeneous microstructures
 - Different material constituents

The microstructure database generation can now be done on **single PC**



Conclusions

- ❑ Rich datasets provide us an opportunity to integrate mechanical and data sciences for rapid prediction, design, and optimization.
- ❑ Data science enables solution of large-scale problems, otherwise not tractable using current methodologies.
- ❑ Reduce Order Models (ROM) such as Principal Component Analysis (PCA), Self-consistent Clustering Analysis (SCA), Multiresolution Clustering Analysis (MCA), help us rapidly generate key datasets.
- ❑ Machine learning techniques such as neural networks (FFNN, CNN, PGNN, etc.) can augment ROMs for extremely fast computations.
- ❑ Combining ROMs with machine learning techniques has the potential to discover, design, and optimize novel complex material systems.
- ❑ Mathematical theories for biological systems are in their infancy; discovery of hypotheses in biological system might be achieved by considering physics, e.g. via a physic guided neural network



Acknowledgement

NORTHWESTERN
UNIVERSITY

Our thanks to...

Funding agencies:

- Air Force Office of Scientific Research
- Army Research Office
- Beijing Institute of Collaborative Innovation
- Bridgestone Corporation
- DMDII (UI Labs)
- Ford Motor Company/National Energy Technology Laboratory
- Lurie Children's Hospital
- NIST CHiMaD I and CHiMaD II
- NIST Gaithersburg
- Northwestern University Data Science Initiative
- NSF
 - GRFP DGE-1324585
 - CMMI MOMS
 - CMMI CPS





References

- [1] Li, H., Kafka, O. L., Gao, J., Yu, C., Nie, Y., Zhang, L., Tajdari, M., Tang, S., Guo, X., Li, G., Tang, S., Cheng, G., & Liu, W. K. (2019). Clustering discretization methods for generation of material performance databases in machine learning and design optimization. *Computational Mechanics*
- [2] Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.
- [3] Zhang, L., Yang, Y., Li H., Gao J., Reno D., Tang S., Liu W.K. Neural network finite element method, in preparation
- [4] Approximation by superpositions of a sigmoidal function, by George Cybenko (1989)
- [5] Multilayer feedforward networks are universal approximators, by Kurt Hornik, Maxwell Stinchcombe, and Halbert White (1989).
- [6] Zhang, L., Yang, Y., Li H., Gao J., Reno D., Tang S., Liu W.K. Neural network finite element method, in preparation
- [7] Watt, J., Borhani, R., & Katsaggelos, A. K. (2016). *Machine learning refined: foundations, algorithms, and applications*. Cambridge University Press.
- [8] Tang, S., Zhang, G., Yang, H., Guo, X., Li, Y., & Liu, W. K. (2019). MAP123: A Data-driven Approach to Use 1D Data for 3D Nonlinear Elastic Materials Modeling, *Computer Methods in Applied Mechanics and Engineering* (Submitted)
- [9] Sigmund, O. A 99 line topology optimization code written in Matlab. (2001). *Structural and multidisciplinary optimization*
- [10] Liu, Z., Bessa, M. A., & Liu, W. K. (2016). *Computer Methods in Applied Mechanics and Engineering*
- [11] Liu, Z., Fleming, M., & Liu, W. K. (2018). *Computer Methods in Applied Mechanics and Engineering*
- [12] Bessa, M. A., Bostanabad, R., Liu, Z., Hu, A., Apley, D. W., Brinson, C., ... & Liu, W. K. (2017). *Computer Methods in Applied Mechanics and Engineering*
- [13] Liu, Z., Kafka, O. L., Yu, C., & Liu, W. K. (2018). In *Advances in Computational Plasticity*
- [14] Tang, S., Zhang, L., & Liu, W. K. (2018). *Computational Mechanics*
- [15] Kafka, O. L., Yu, C., Shakoor, M., Liu, Z., Wagner, G. J., & Liu, W. K. (2018). *JOM*
- [16] Shakoor, M., Kafka, O. L., Yu, C., & Liu, W. K. (2018). *Computational Mechanics*
- [17] Li, H., Kafka, O. L., Gao, J., Yu, C., Nie, Y., Zhang, L., ... & Tang, S. (2019). *Computational Mechanics*
- [18] Zhang, L., Tang, S., Yu, C., Zhu, X., & Liu, W. K. (2019). *Computational Mechanics*
- [19] Gao, J., Shakoor, M., Jinnai, H., Kadowaki, H. Seta, E., Liu, W. K. An Inverse Modeling Approach for Predicting Filled Rubber Performance. (2019) *Computer Methods in Applied Mechanics and Engineering*