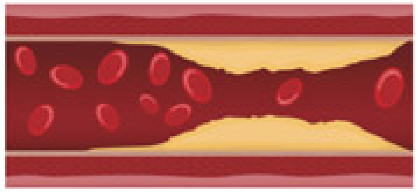
Coronary Artery Disease (CAD) and its diagnosis

- CAD caused approximately 370,000 fatalities in the US every year.
- CAD is usually caused by the buildup of cholesterol and fatty deposits (called plaques) inside the arteries, which will limit or stop blood flow to the heart.

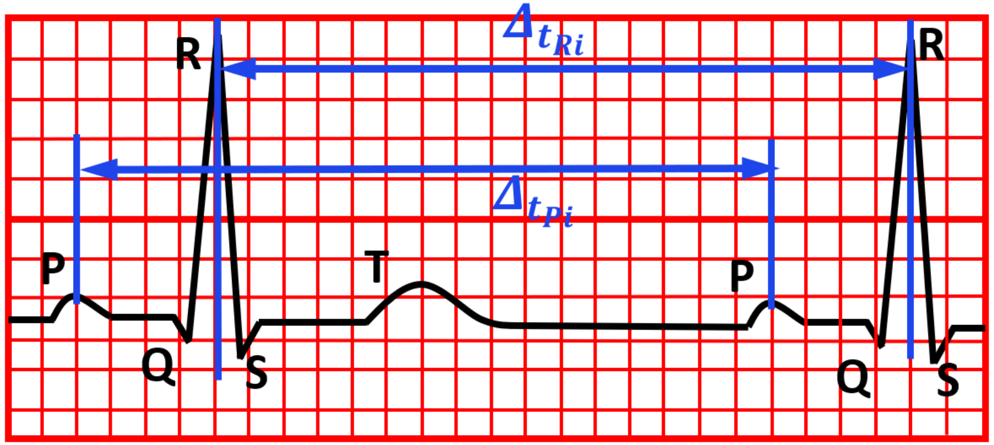


Partially blocked artery

Fully blocked artery

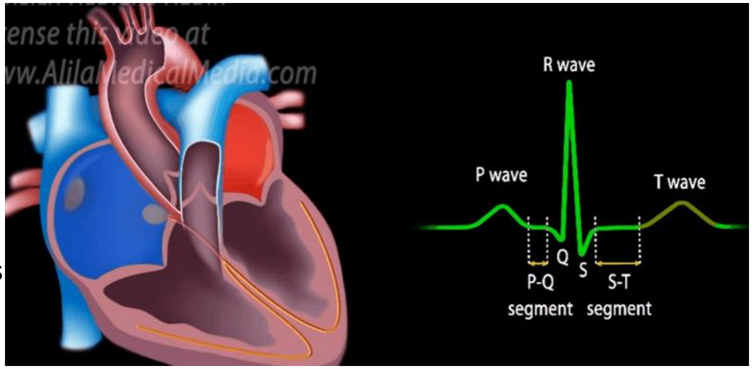
[1] https://www.cdc.gov/heartdisease/facts.htm
[2] Coronary Artery Disease. (n.d.). Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/16898-coronary-artery-disease
[3] https://www.cvphysiology.com/Arrhythmias/A009
[4] https://www.alilamedicalmedia.com/

- CAD can be diagnosed using electrocardiogram (ECG), which checks the heart's electrical activity. But the accuracy of diagnosis is highly dependent on physician's training.



Common ECG waveform

- **Each cycle** of ECG contains **P, Q, R, S, T** waves
- **P wave**: atrial depolarization (atrial contraction)
- **QRS complex**: ventricle depolarization (ventricle contraction)
- **T wave**: ventricle polarization (ventricle relaxation)
- $\Delta_{t_{Ri}}$: *i*-th time interval of 2 consecutive R wave peaks
- $\Delta_{t_{Pi}}$: *i*-th time interval of 2 consecutive P wave peaks

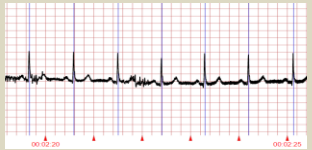


Coronary Artery Disease (CAD) diagnosis using ElectroCardioGram (ECG)

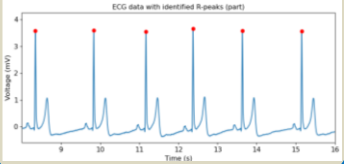
Knowledge Database

Data preparation

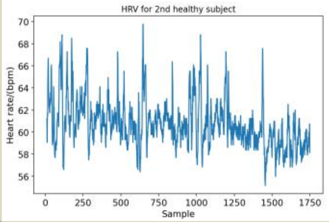
Raw ECG file from PhysioNet [2]



Automatic R peak detection [3]

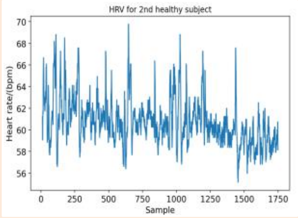


Heart rate variability (HRV)

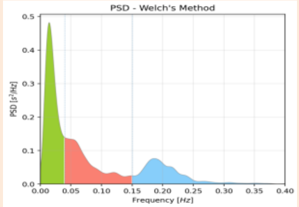


Mechanistic feature extraction

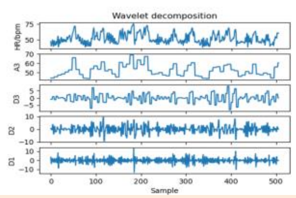
Time domain (mean， std)



Freq domain (powers)



Time-freq domain (entropies)



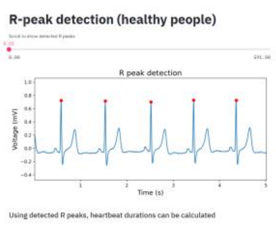
Mechanistic learning through classification

Classification of **patients** and **healthy people** through support vector machine **(SVM)**

- **SVM** finds an **optimal hyperplane** that has **largest distance** to **support vectors** (nearest data points)
- **SVM** works well for problems with **small datasets**

System and design

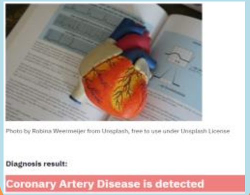
A **web-based** interactive **mini-app** for CAD diagnosis



Optimized system

Uploaded raw ECG signal

CAD diagnosis result



R peak detection is widely used to diagnose heart rhythm irregularities [2]

[1] www.mayoclinic.org/tests-procedures/ekg
[2] https://physionet.org/
[3] Lourenço et. al., Biosignals, 2012, pp. 49-54

Data Preparation

- Data balance

① Fantasia database
  - **40 healthy subjects**
  - **Duration: 120 min**
  - **Frequency: 250 Hz**
  - 1 signal for each

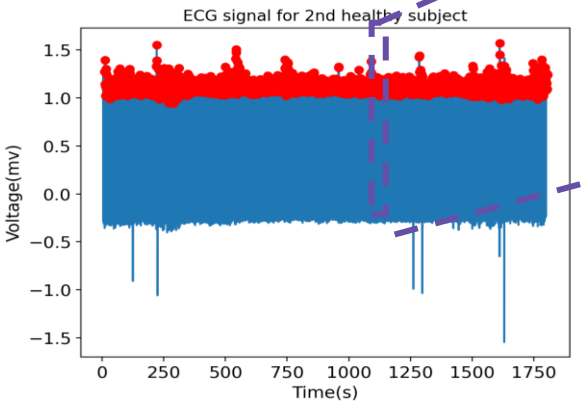
② St Petersburg database
  - **7 CAD patients**
  - **Duration: 30 min**
  - **Frequency: 257 Hz**
  - 12 signals for each

https://physionet.org/about/database/

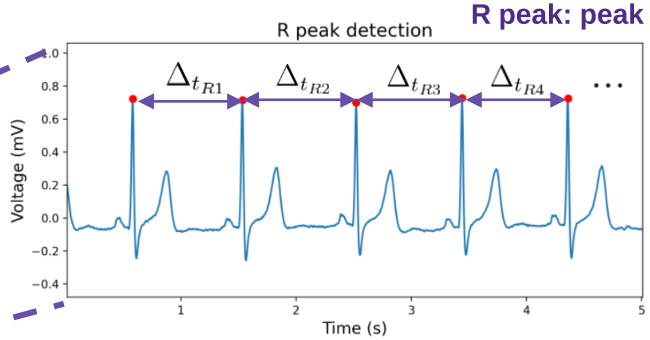
- **Segmentation**
- **Upsampling**

- ECG denoising: noises will affect the detection of R peaks. Low and high frequency noises are filtered using the Butterworth filter
- Automatic R peak detection and heart rate variability (HRV) derivation

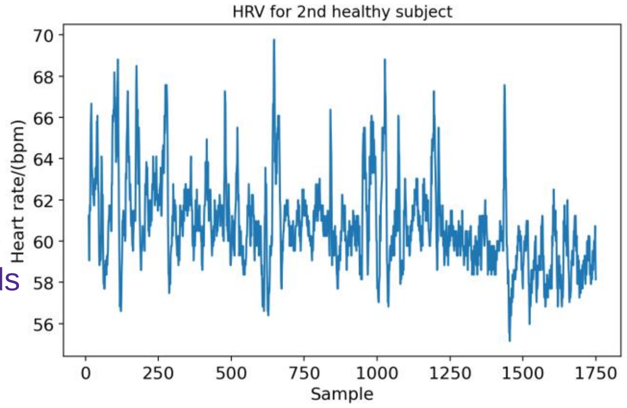
ECG signal



R peak: peak of the R wave



Derived HRV



$$HRV_i = \frac{60}{\Delta_{tRi}}$$

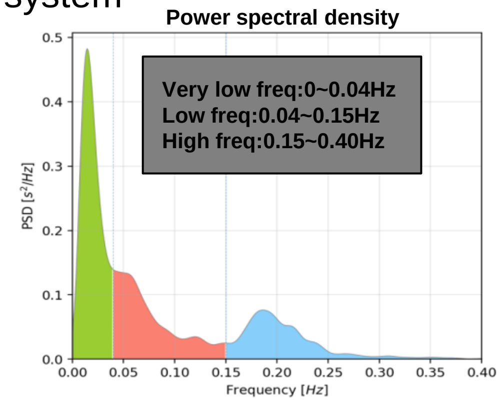
60 seconds

i-th heart rate

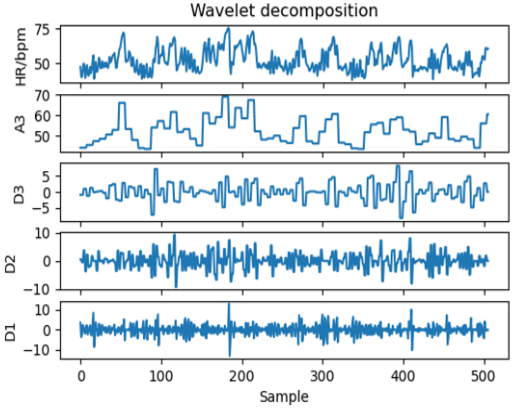
i-th time interval of 2 consecutive R wave peaks

Feature Extraction

- Time domain (3)-linear statistical quantities
  - **Mean** of heartbeat durations $\overline{\Delta_t} = \sum_{j=1}^{n} \Delta_{t_i}$
  - **Standard deviation** of heartbeat durations $SD = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(\Delta_{t_j} - \overline{\Delta_t}\right)^2}$
  - **Standard deviation** of heartbeat duration **differences** $SDSD = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(\Delta\left(\Delta_{t_j}\right) - \overline{\Delta(\Delta_t)}\right)^2}$
- Frequency domain (4)-reveals how heart rate is controlled by the nervous system
  - **Total power**: detecting abnormal autonomic activity
  - **Low frequency power**: sympathetic modulation
  - **High frequency power**: parasympathetic modulation
  - **LF/HF ratio**: sympathetic/parasympathetic balance

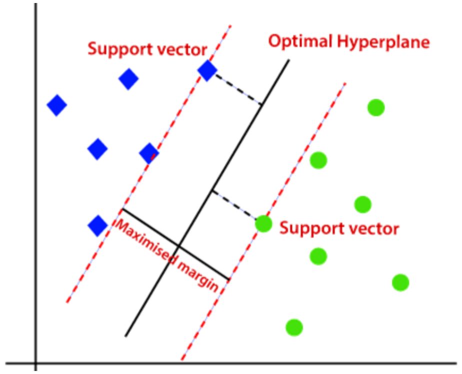


- Time-frequency domain (12)-nonlinear analysis of the signal complexity
  - **Shannon entropy:** measure of data uncertainty and variability
  - **Approximation entropy:** quantify the amount of regularity and unpredictability of fluctuations
  - **Sampling entropy:** assess complexities of physiological signals



Classification using Support Vector Machine (SVM)

- **Scale input features** before classification
- Label the outputs for **healthy** subjects as **1**; **CAD** patients as **0**
- **SVM** intends to find an **optimal hyperplane** that has **largest distance** to **support vectors**
- **SVM** is good for **small dataset**; **linear kernel** is used in the project



https://www.analyticsvidhya.com/blog/2021/05/5-classification-algorithms-you-should-know-introductory-guide/

```python
from sklearn import svm
scaler = preprocessing.StandardScaler().fit(X)
joblib.dump(scaler, 'data_scaler.pkl')
X = scaler.transform(X)
clf = svm.SVC(kernel='linear', C=1000, random_state=0)
```

Linear SVM model

- **5-fold cross validation** is used to assess the performance of the ML model

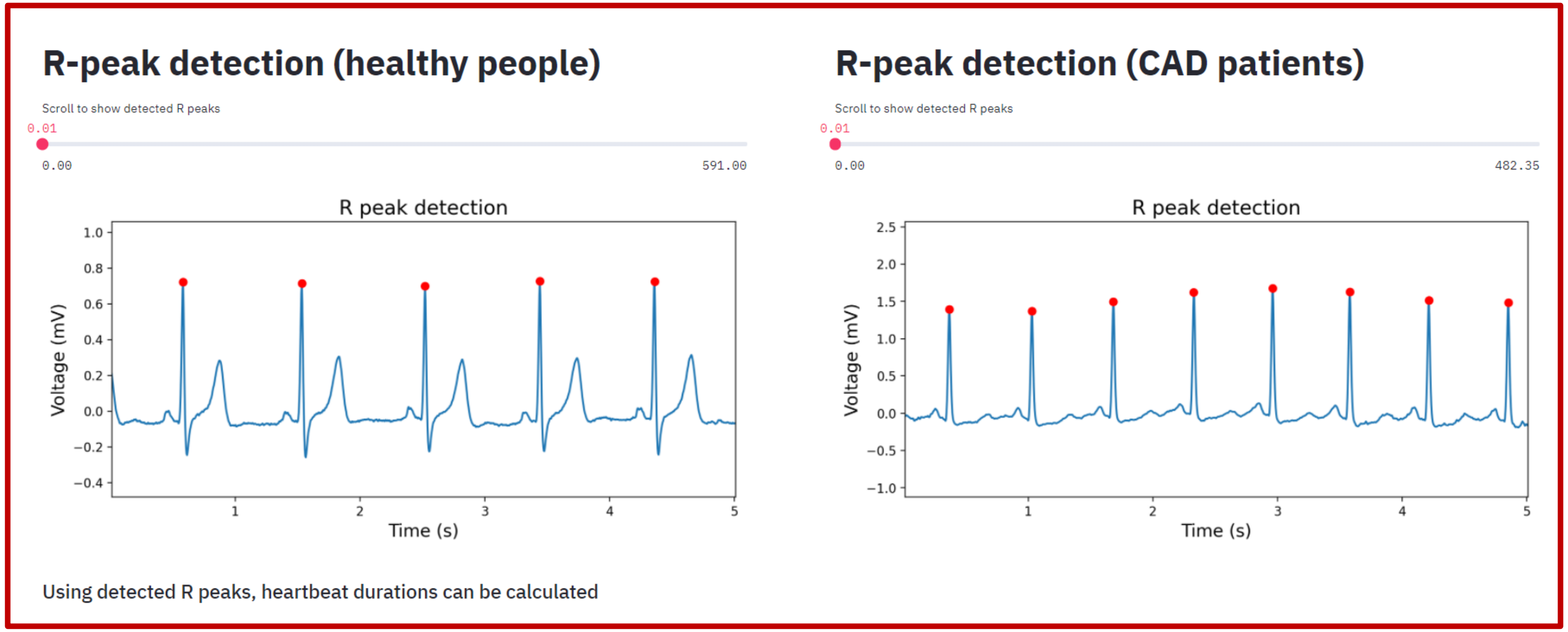
```
In [20]: runfile('F:/MDS/svmTraining.py', wdir='F:/MDS')
Fold:  1, Accuracy: 0.962, F1 score 0.961
Fold:  2, Accuracy: 0.949, F1 score 0.947
Fold:  3, Accuracy: 0.975, F1 score 0.974
Fold:  4, Accuracy: 0.987, F1 score 0.987
Fold:  5, Accuracy: 0.910, F1 score 0.914

Cross-Validation accuracy: 0.957 +/- 0.026
Cross-Validation F1 score: 0.957 +/- 0.025
```

Training results

System and Design (facilitating CAD diagnosis)

- **Inputs: raw ECG signals (csv file)** from a standard ECG recording
- **Outputs: diagnosis result** (**healthy** or **CAD**)
- A **web-based interactive mini-app** which facilitate the physicians with the CAD diagnosis process is developed



Mini-app interface

Computational Materials Science 214 (2022) 111703

Contents lists available at ScienceDirect

Computational Materials Science

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

ELSEVIER

Full Length Article

Knowledge database creation for design of polymer matrix composite

Hannah Huang [a,1], Satyajit Mojumder [b,1], Derick Suarez [a], Abdullah Al Amin [a], Mark Fleming [a], Wing Kam Liu [a,*]

[a] *Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA*
[b] *Theoretical and Applied Mechanics Program, Northwestern University, Evanston, IL 60208, USA*

A R T I C L E I N F O

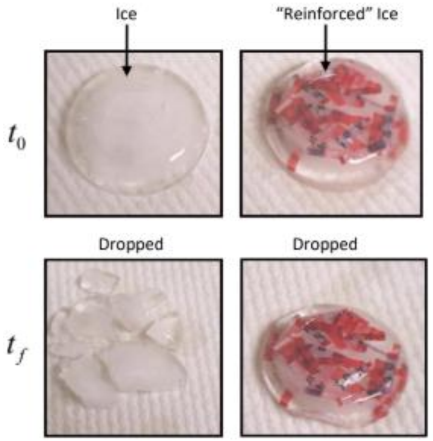
Keywords:
Mechanistic data science
Unidirectional fiber
Polymer composite
Materials design
Mechanistic features
Dimension reduction

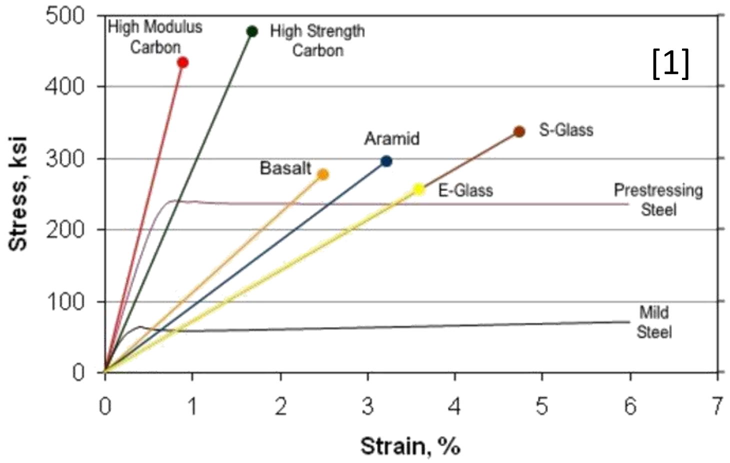
A B S T R A C T

We present a mechanistic data science (MDS) framework capable of building a composite knowledge database for composite materials design. The MDS framework systematically leverages data science to extract mechanistic knowledge from composite materials system. The composite response database is first generated for three matrix and four fiber combinations using a physics-based mechanistic reduced-order model. Next, the mechanistic features of the composites are identified by mechanistically analyzing the composites stress–strain responses. A relationship between the composite properties and the constituents' material features are established through a mechanics constrained data science-based learning process after representing materials in latent space, following a dimension reduction technique. We demonstrate the capability of predicting a composite materials system for target properties (material elastic properties, yield strength, resilience, toughness, and density) from the MDS created knowledge database. The MDS model is predictive with reasonable accuracy, and capable of identifying the materials system along with the tuning required to achieve desired composite properties. Development of such MDS framework can be exploited for fast materials system design, creating new opportunity for performance guided materials design.

What is a Composite Material?

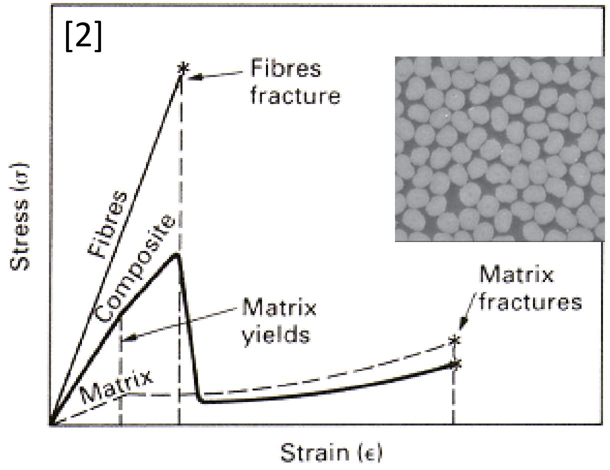
Let's make a simple composite material. Reinforcing ice (the **matrix**) with newspaper (the **reinforcement**) improved the material's resistance to fracture (**toughness**).





One type of composite, **fiber-reinforced** composites, combines the high **modulus** of fibers (such as carbon and glass) with other materials.

- In general, fibers are quite brittle compared to matrix
- In **fiber-reinforced polymer** composites, the polymer matrix acts as a binder for the fiber
- The **matrix** imparts toughness into the material
- We can achieve **lightweight** and **high strength** materials with **lower densities** than metals



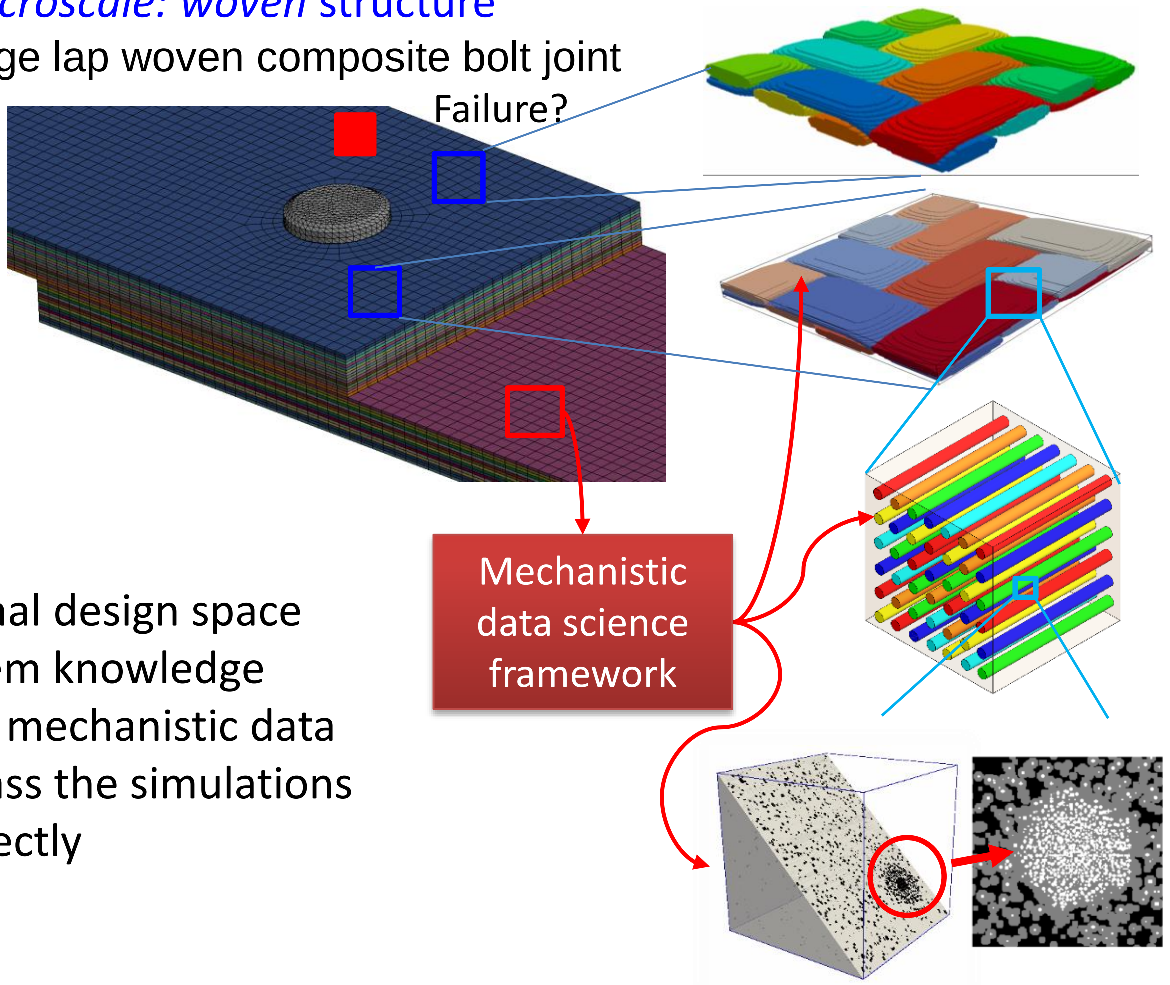
[1] https://www.abbottaerospace.com/aa-sb-001/4-materials/4-1-composite-materials/4-1-3-fundamental-behavior-of-carbon-fiber-epoxy-resin-composite-laminates/
[2] https://slideplayer.com/slide/4235571/

Design of polymer matrix composites at multiple scales

Macroscale: woven structure

Singe lap woven composite bolt joint

Location dependent microstructure design



Design of interphase and agglomeration

- High dimensional design space
- Building a system knowledge database using mechanistic data science to bypass the simulations and predict directly

Design variables

Mesoscale: woven structure

- Fiber volume fraction
- Yarn angle
- Yarn geometry
- Matrix and fiber choice

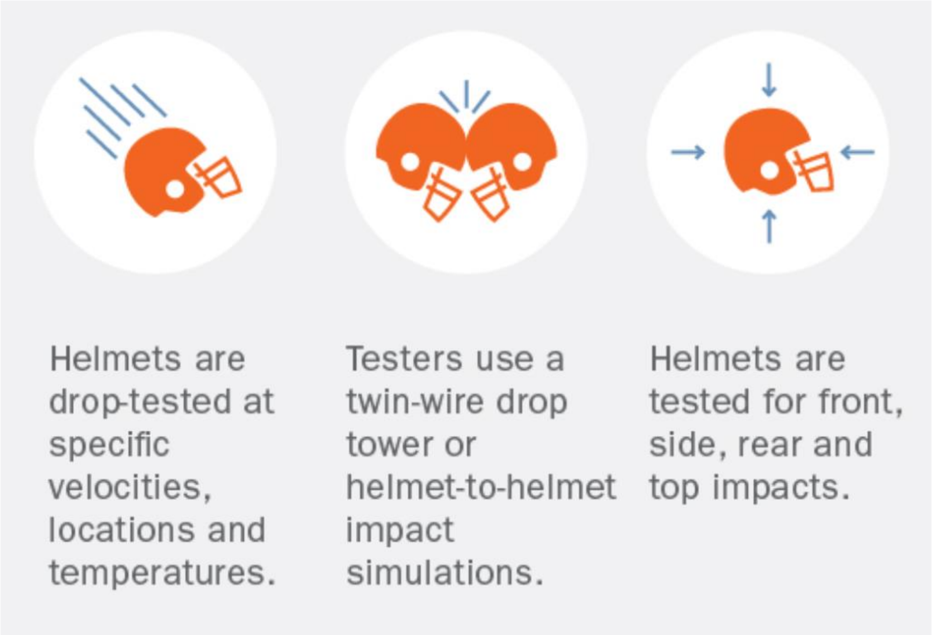
Microscale: Unidirectional structure

- Fiber volume fraction
- Matrix and fiber choice
- Fiber geometry

Nanoscale: reinforced matrix

- Filler volume fraction
- No of agglomerations
- Interphase size and properties
- Filler and matrix choice

Design of a material system for composite football helmet



Design parameters:
- Polymer and fiber choice
- Fiber volume fraction , etc..

Design criteria:
- Lightweight (low density)
- High strength
- High toughness (absorb impact energy)

Why do we need a better helmet?
SAFETY!

Helmets are crucial to prevent brain injuries.

From 2012-2019, an average of 242 NFL players per year sustained concussions. At the high school level, football is responsible for 60% of all sports related concussions. [11]

[11] https://www.thespeedycheetah.com/products/northwestern-wildcats-xp-replica?dfw_tracker=2623-922099945&gclid=Cj0KCQjw6575BRCQARIsAMp-ksMqcLC0Lb20NX3yj8f6ecTFEc9DGoA4qvgE3icNgOWkeoTzupedq6EaAo6KEALw_wcB
[12] https://www.cognitivefxusa.com/blog/football-concussion-prevention-and-recovery

System and Design Problem

System Description:

Composite materials are needed for an array of applications (football helmet for example). Here we consider unidirectional fiber polymer composites, where the main features include volume fraction, fiber radius, and matrix and fiber material properties.

Modeling Objectives:

Develop a mechanistic data science model that predicts a quantifiable output (such as the stress response, toughness, damage, etc.) based on input features.

https://www.sciencemag.org/news/2017/07/ninety-nine-percent-ailing-nfl-player-brains-show-hallmarks-neurodegenerative-disease

Design Objective:

Design of new composite materials system for desired performance in application. (Ex. Higher toughness of Helmet materials)

Six steps of mechanistic data science in composite materials

Design for absorbing impact energy

Surrogate models

Neural network model

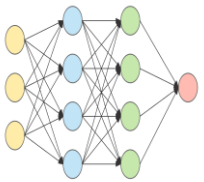
PCA model

Prediction of materials system

Microstructure property

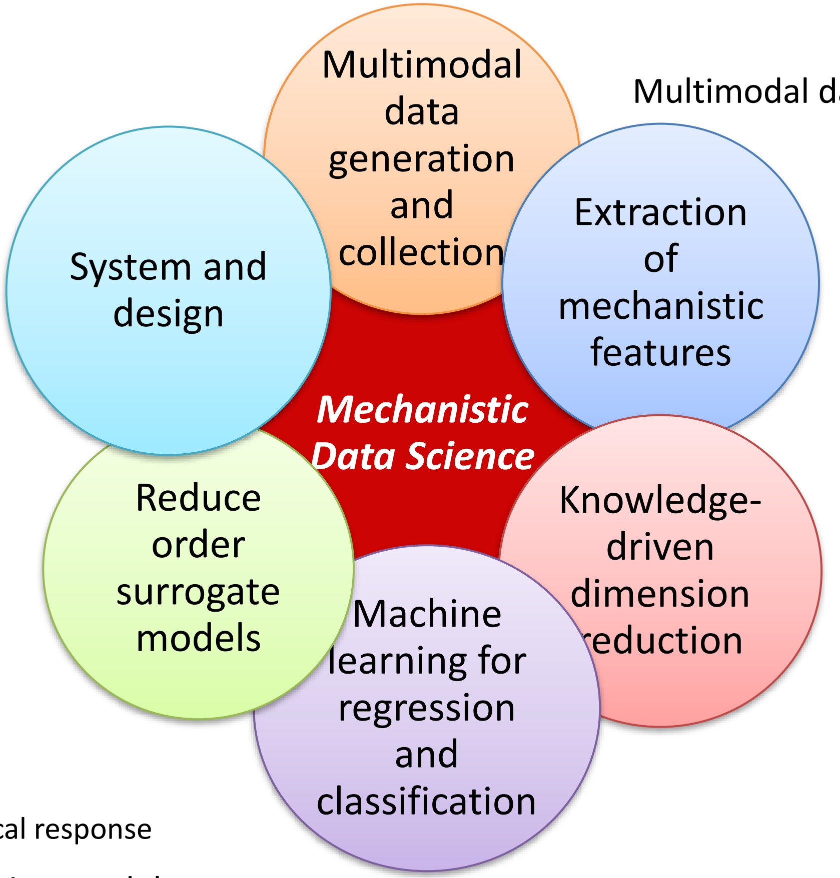
Materials property

Physical property

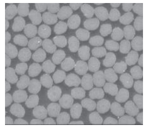


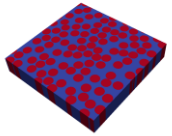
Mechanical response

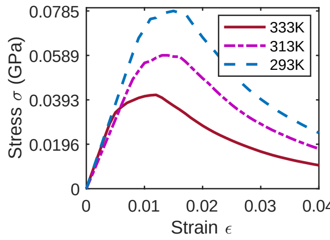
Regression model



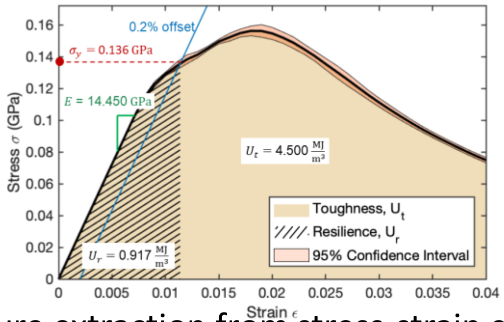
Possible materials: short fiber, UD, foam



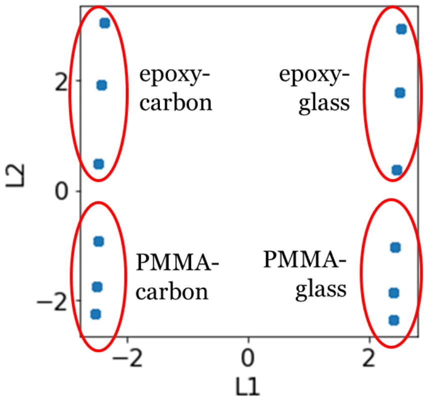




Multimodal data from simulation and experiment



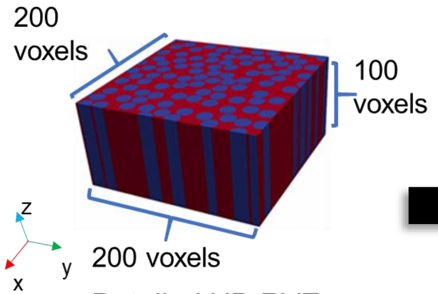
Feature extraction from stress-strain curve



Clustering in latent variable space

Data generation through tensile simulations

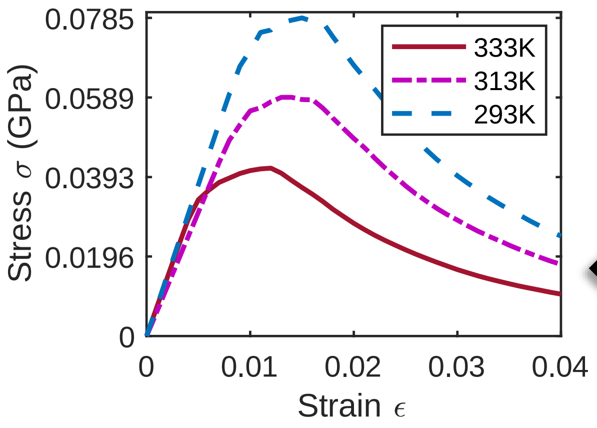
Model input



Detailed UD RVE (4×10^6 voxels)

Matrix and filler properties

Model output



Approximately 1.5 mins for the online simulation

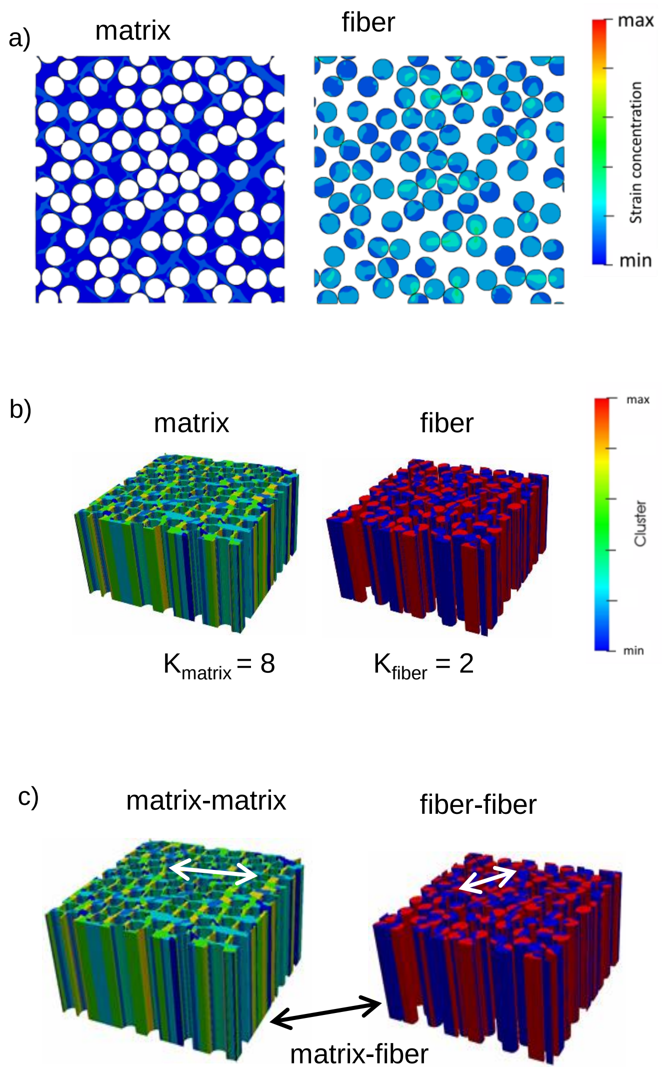
Self-consistent Clustering Analysis

Offline

a) Data collection (local strain response)
b) Data compression by unsupervised learning (K-means clustering)
c) Interaction tensor calculation for clusters

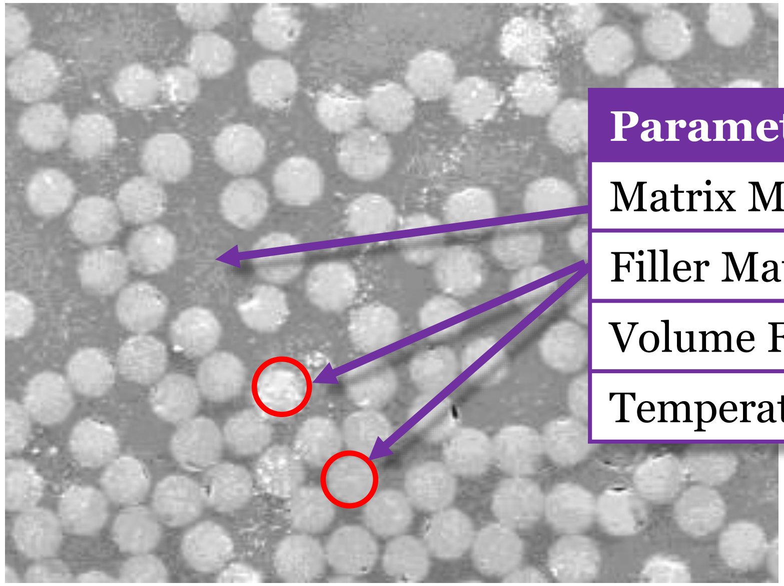
Online

Solve Lipmann-Schwinger equations for the clusters to predict stress-strain response



Z. Liu et al. / Comput. Methods Appl. Mech. Engrg. 306 (2016) 319–341

Design of experiment and database generation



Design of Experiment Summary

Parameter	Range
Matrix Material	Epoxy, PMMA,PET
Filler Material	Carbon , Glass, Kevlar fiber
Volume Fraction	[0.01, 0.50] (50 samples)
Temperature	[213, 333] (K) (3 samples)

*5 realizations for each microstructure descriptors

$$volume\ fraction = \frac{filler\ volume}{total\ volume}$$

Input material properties for simulations

Matrix Material
- density
- elastic modulus
- Poisson's ratio
- yield strength
- hardening parameter

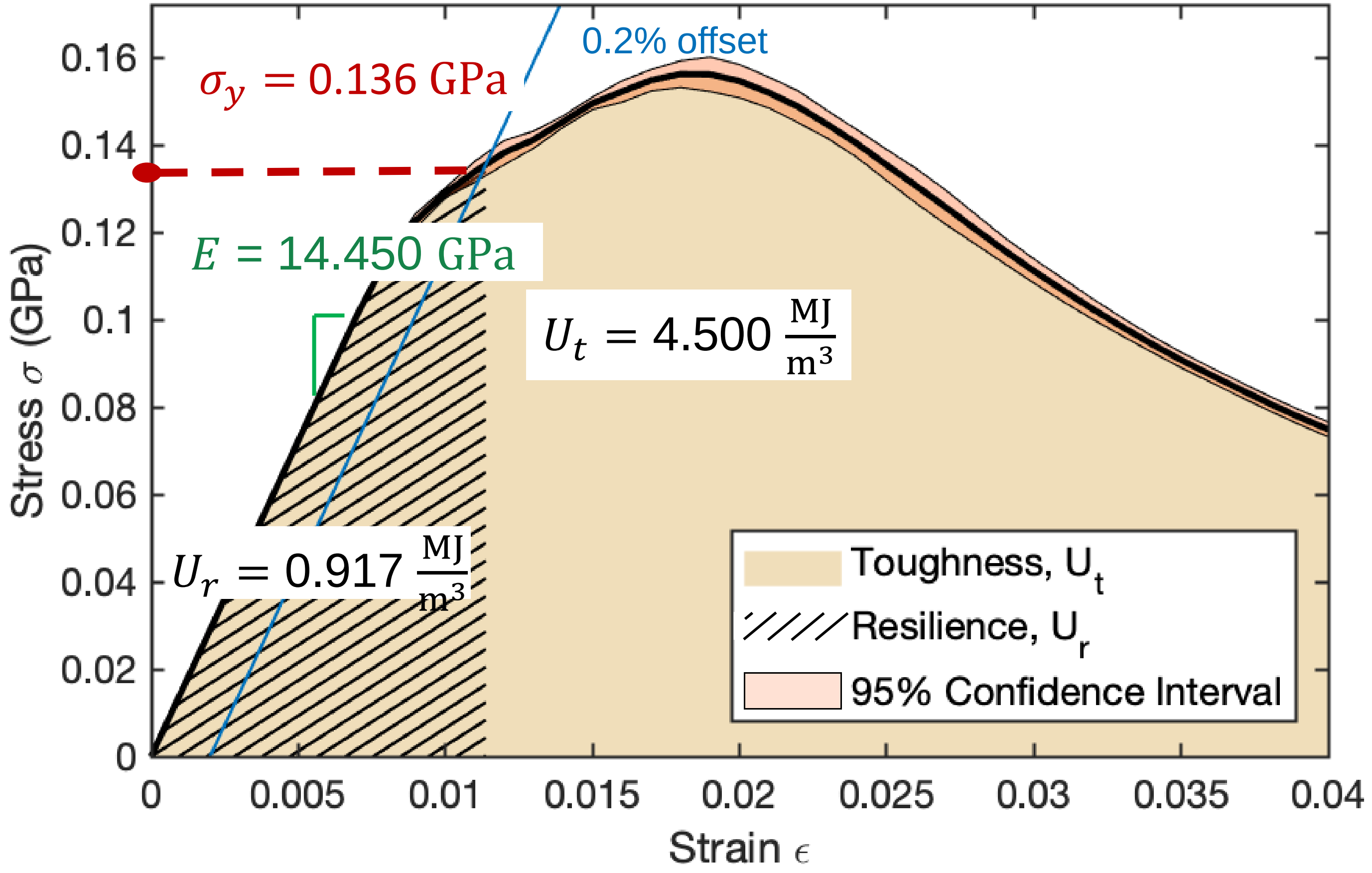
Filler Material
- density
- x-elastic modulus
- y-elastic modulus
- z-elastic modulus
- x-Poisson's ratio
- y-Poisson's ratio
- z-Poisson's ratio
- x-shear modulus
- y-shear modulus
- z-shear modulus

Total 6500 simulations are performed to build the database

Tribological behaviour of unidirectional carbon fibre-reinforced epoxy composites - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Cross-sectional-view-of-unidirectional-carbon-fibre-reinforced-epoxy-composites-taken-by_fig1_314138488 [accessed 21 Aug, 2020]

Mechanistic feature extractions from stress-strain curves

Simulated stress-strain curve of 0.25 volume fraction of Carbon in Epoxy at 213K

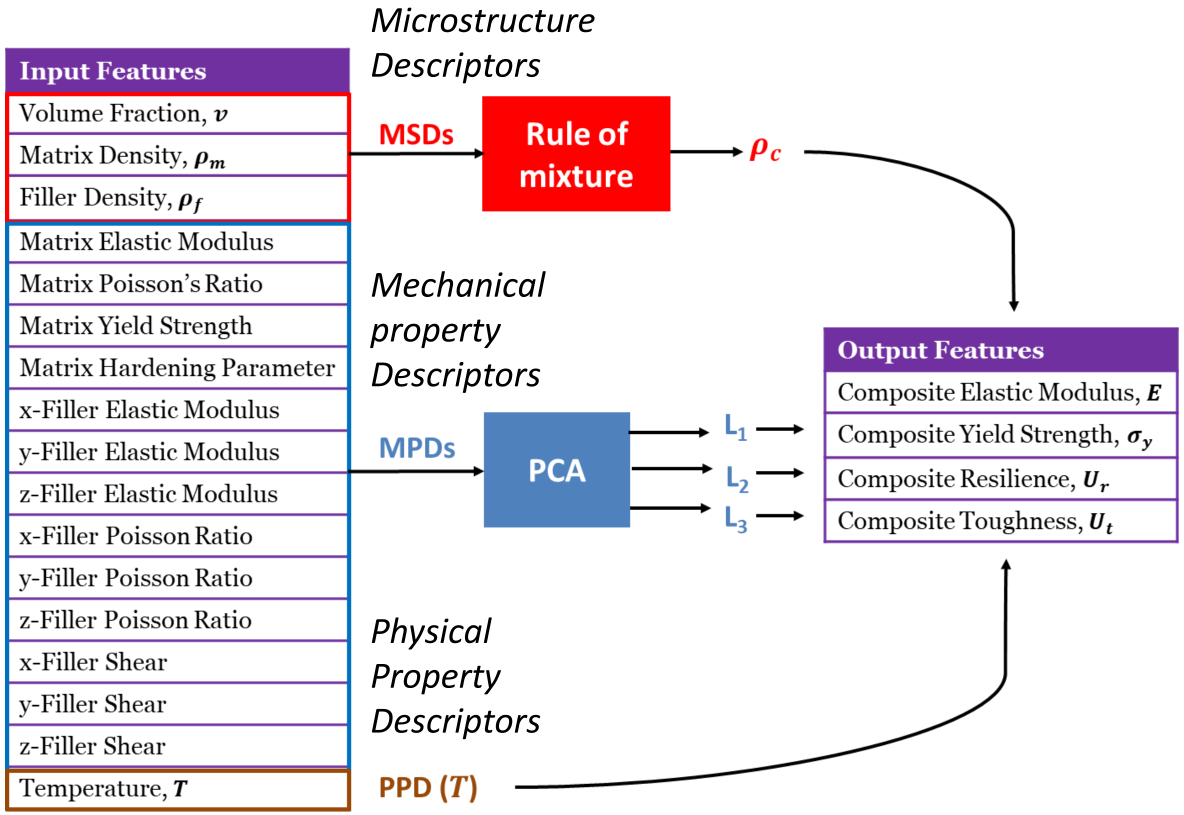


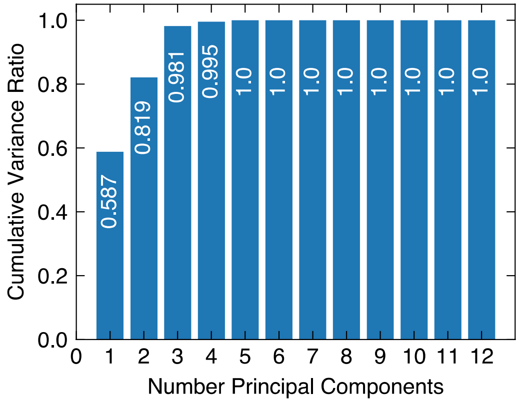
Mechanistic features:

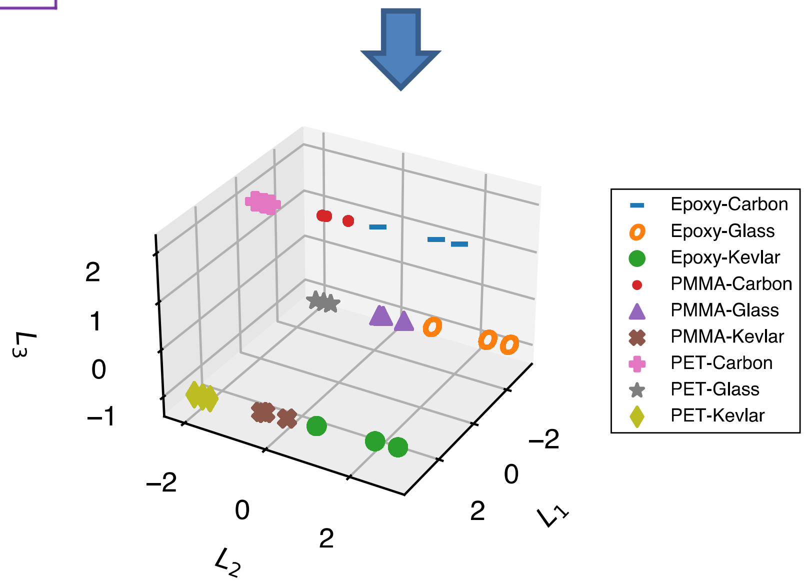
- Elastic modulus
- Yield strength
- Toughness
- Resilience

These mechanistic features are a function of the composite microstructures and materials properties.

Reducing problem dimension through mechanistic relation and PCA

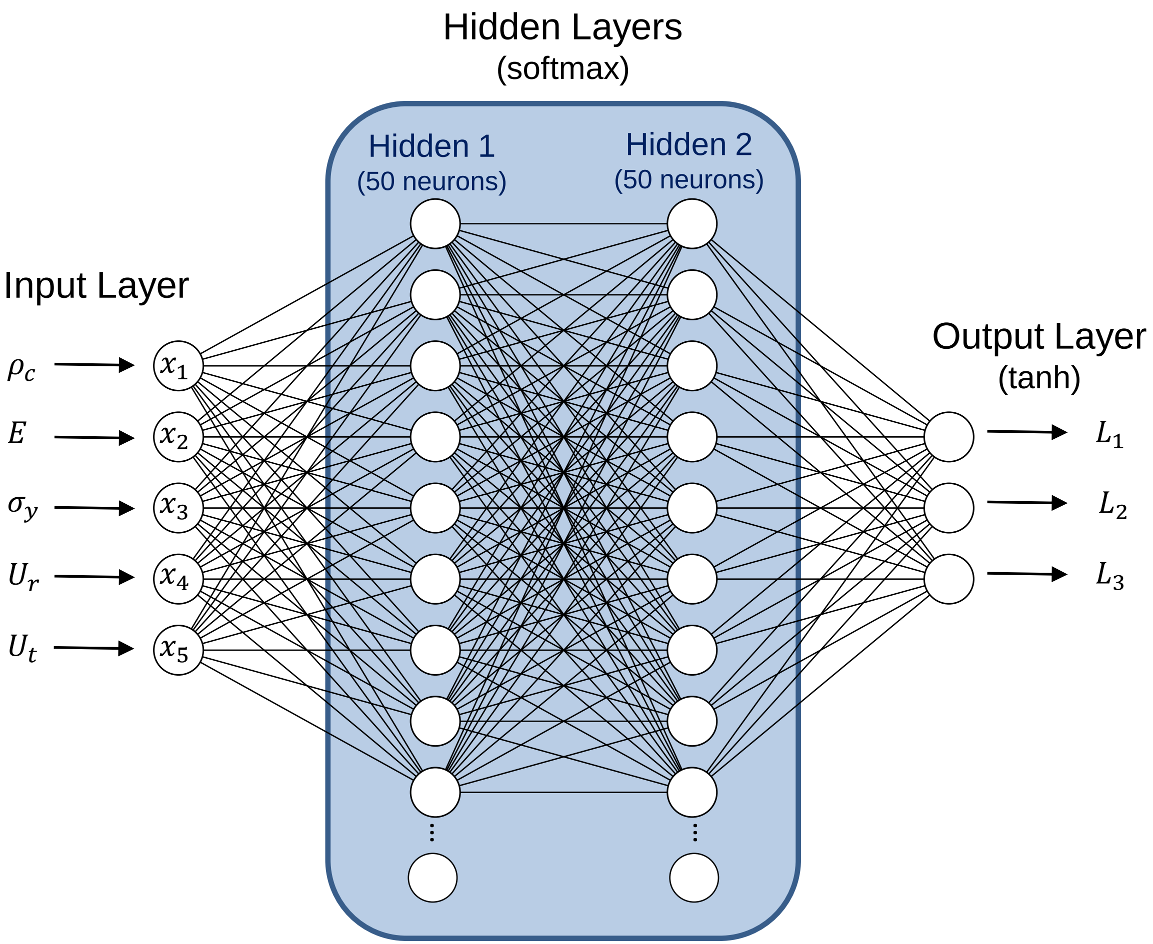


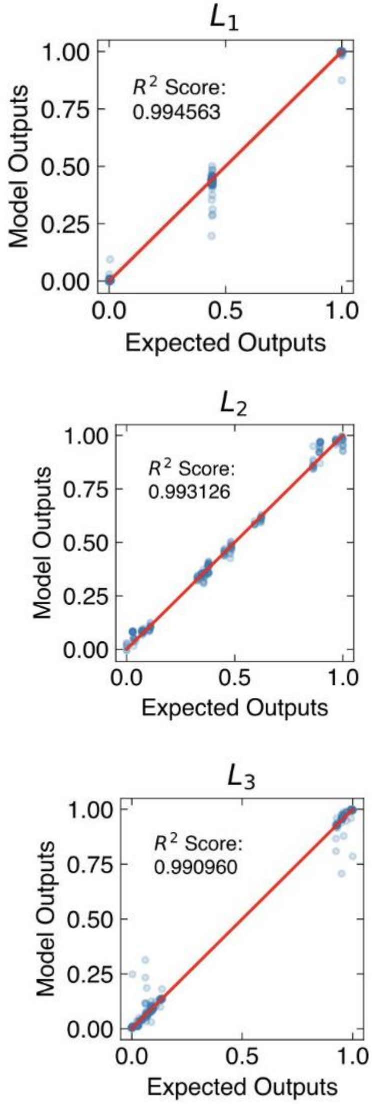




- Reducing high dimensional (14 dimensions) materials space to latent materials space
- Materials form clusters in the latent space
- Explore the latent space for new materials system

Neural network-based regression model identifies mechanistic features relation with latent property space





- Very high accuracy in the trained model

Using MDS model for the prediction of new materials system of desired performance

Design specifications

ρ_c
E
σ_y
U_r
U_t

Design decision from MDS

- Materials system
- Microstructure details
- Operating condition
- Desired toughness

System knowledge database **built using mechanistic data science**

Trained neural Network from regression step

Materials properties dimension reduction using Principal component analysis

Stress-strain database generated using SCA from data generation step

Neural network prediction

$L_1^*, L_2^*, \dots, L_n^*$

Materials representation latent materials property space

Inverse PCA

Microstructure, matrix and fillers identification

Compare with Closest material system

Identify material system and suggest possible modifications

MDS guides the materials design process for user defined properties

Properties	Expected	Predicted (MDS)	Difference (Percentage)
Matrix Elastic Modulus (MPa)	2250	2070.435	7.9806
Matrix Poisson's Ratio	0.34	0.343256	0.9576
Matrix Yield Strength (MPa)	21.64725	20.03131	7.4649
Filler Elastic Modulus 1 (MPa)	118000	116153.9	1.5645
Filler Elastic Modulus 2 (MPa)	7200	7289.216	1.2391
Filler Elastic Modulus 3 (MPa)	7200	7289.216	1.2391
Filler Poisson 1	0.27	0.268592	0.5213
Filler Poisson 2	0.27	0.268592	0.5213
Filler Poisson 3	0.34	0.339824	0.0517
Filler Shear 1 (MPa)	2800	2802.993	0.1069
Filler Shear 2 (MPa)	2800	2802.993	0.1069
Filler Shear 3 (MPa)	2700	2733.786	1.2513
Total Mean Relative Difference:			1.9171

Features	Value
ρ_c	1477.5 Kg/m^3
E	14644.76 MPa
σ_y	140.84 MPa
U_r	0.8972 MJ/m^3
U_t	4.3084 MJ/m^3

Desired properties

Material system 1: Epoxy/Carbon
Temperature: 230K
Volume fraction of filler: 0.35
Material system 2: PMMA/Carbon
Temperature: 293K
Volume fraction of filler: 0.01
Material system 3: PET/Carbon
Temperature: 300K
Volume fraction of filler: 0.35

Design recommendation: *Add nanofiller in the matrix materials to augment stiffness*