Dataset Overview

- Synthetic dataset
- Each sample:
  - `eps`, global (laminate level) strain vector: $\boldsymbol{\varepsilon} \in \mathbb{R}^6$
  - `plies`, dictionary with 4 orientations: $p \in \{0, \pm 45, 90\}$
  - Per ply, 4 failure indices corresponding to 4 different cases $c \in \{ft, fc, mt, mc\}$:

	Tension (t)	**Compression (c)**
Fibre (f)	F_{ft}	F_{fc}
Matrix (m)	F_{mt}	F_{mc}

- Total of 8/16 targets (depending if tension/compression are mutually exclusive)

Task 1: Regression of Maximum Failure Index

Goal: Predict only the largest failure index:

$$y = \max_{p,c} F_{p,c}$$

Input: $\boldsymbol{\varepsilon} \in \mathbb{R}^6$
Output: $\hat{y}_{max} \in \mathbb{R}^+$

Models:

- Linear / Ridge Regression
- Random Forest, Gradient Boosting (XGBoost)
- Feed-forward Neural Network (MLP)

Task 2: Multi-Output Regression

Goal: Predict all ply-level failure indices from global strain.

$$y = \{F_{p,c}\}_{p \in \{0,\pm 45,90\},\ c \in \{ft,fc,mt,mc\}}$$

Input: $\varepsilon \in \mathbb{R}^6$
Output: $\hat{y} \in \mathbb{R}^{8 \text{ or } 16}$

Models:

- Multi-output Linear / Ridge Regression
- Random Forest, Gradient Boosting (XGBoost)
- Feed-forward Neural Network (MLP)

Task 3: Binary Classification (Failure Detection)

Goal: Classify whether the structure is in a safe or failed state.

$$\text{Label} = \begin{cases} \text{“failure”}, & \max\limits_{p,c} F_{p,c} > 1 \\ \text{“no failure”}, & \text{otherwise} \end{cases}$$

Input: $\varepsilon \in \mathbb{R}^6$
Output: Binary class label

Models:

- Logistic Regression
- Random Forest / XGBoost
- Support Vector Machine
- Neural Network Classifier

Task 4: Multi-Class Classification (First Failing Ply or Mode)

Goal: Identify which ply–mode pair fails first, or predict a safe state.

$$(p^*, c^*) = \begin{cases} \arg\max\limits_{p,c} F_{p,c}, & \text{if } \max\limits_{p,c} F_{p,c} > 1 \\ \text{“no failure”}, & \text{otherwise} \end{cases}$$

Possible labels:

- 16 classes (4 plies $\times$ 4 modes) + 1 non-failure state

Models:

- Random Forest, XGBoost
- Logistic Regression (multi-class)
- SVM
- Neural Network Classifier