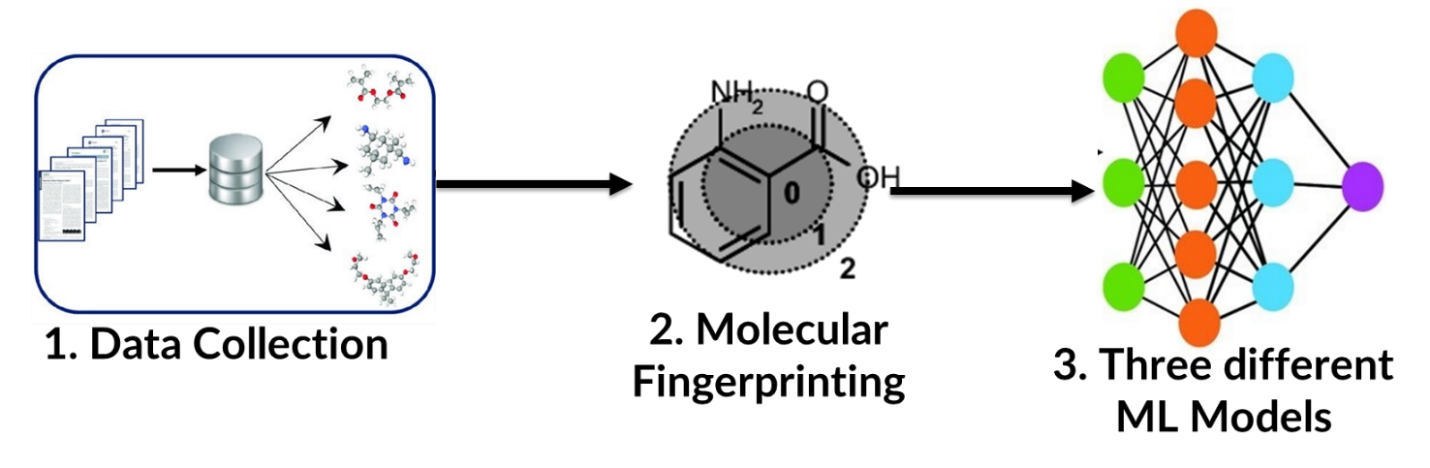
# Machine Learning Tailored Dynamic Mechanical Analysis (DMA) curve for Shape Memory Polymers.

## Abstract

## Introduction

## Methods

In this section, we present the procedures of our work. This includes the establishment of four datasets, the establishment of four machine learning models, and the prediction of the TSMP properties. Figure 1 shows the stages of this work.

The detailed description of the outline is as follows:

1. Establishment of the datasets
   1. Data was sourced from literature considering only pure TSMPs i.e. TSMPs without any form of reinforcement
   2. Particularly, the structures of the monomers the TSMPs are made of, their respective molecular weight, and the four properties of interest were considered for this work.
2. Molecular fingerprinting
   1. The structures of the monomers were canonicalized by adopting the Simplified Molecular Input Linear Entry System (SMILES) ; a linear notation system.
   2. The canonicalized monomers were then fingerprinted into high dimensional vectors using the Morgan Fingerprinting, a circular molecular fingerprinting system.
3. Establishment of the ML models
4. Three non linear ML models were selected for this work for comparison sake.
5. The models namely Gaussian Process Regression (GPR), Support Vector Regression (SVR), and Artificial Neural Network (ANN) were compared based on three metrics.
6. Although the main metric of focus was the mean absolute percentage error (MAPE), two other metrics, mean squared error (MSE), and mean absolute error (MAE) were added to further confirm the best performing model.

### SMILES

SMILES is an abbreviation for Simplified Molecular Input Linear Entry System. This is a linear notation widely used in represent structure of molecules into a linear form. SMILES uses a very short and natural grammar, enables rigorous structural specification based on the ideas of molecular graph theory, and is well-suited for high-speed machine processing [1]. SMILES creates a simplified version of molecular structures by changing them into graphs. Although several other notations such as the Wiswesser notation and the Representation of Organic Structures Description Arranged Linearly (ROSDAL notation) exist, the SMILES notation is more lax and easy to learn [1,2]. Other methods of converting molecular structures to strings are more complicated and involving. In chemical informatics, other variations of the SMILES also exist such as DeepSMILES[ref] and BigSMILES[ref].

### Molecular Fingerprinting

### GPR

### ANN

### SVR

### MAE,MSE, MAPE

## Results and Discussion

The glassy modulus, storage modulus and the temperatures at which they occur were studied. Four datasets for each of the four properties of the DMA curve i.e. glassy and rubbery moduli and their respective temperatures. From various references, about an average of 196 datapoints was collected for each of the datasets. Collected from these references were the properties of interest, the structures of the respective monomers and their molar ratios. In this work, only pure shape memory polymers were considered. Also, even if there were catalyst in the synthesis of these monomers, they were taken into account.

For each dataset, the structures of the amine and epoxy were converted in the Simplified Molecular Linear Input Line Entry System (SMILES) notation. SMILES represents the molecular structures as graphs.

## Conclusion

## Experimental Section