OVERCOMING MODELING CHALLENGES IN THERMOSET SHAPE MEMORY POLYMERS: A MACHINE LEARNING APPROACH

“machine learning tailored Dynamic Mechanical Analysis (DMA) curve ”.

# Abstract

[insert abstract here]

# Introduction

Shape Memory Polymers (SMPs) are gradually becoming popular in today’s material research. SMPs are smart stimuli-responsive polymeric materials that respond to external stimuli including temperature, light irradiation, solvents, electrical current, magnetic field, and electromagnetic field to assume a memorized permanent shape that may be programmed [1]. Due to their self-healing capabilities, shape memory polymers are widely used in aviation, medicine, fashion, and structural models. Shape-memory polymers. Due to their self-healing capabilities, shape memory polymers are widely used in aviation, medicine, fashion, and structural models [2]. Predicting the thermomechanical properties of thermoset shape memory polymers (TSMPs) is crucial for their effective design and application. However, the complex topological nature of TSMPs presents a significant challenge in modeling these properties, often resulting in time-consuming processes.

The role of machine learning (ML) in the discovery and synthesis of new functional organic molecules with specialized applications cannot be overstated. Its remarkable capacity to identify and analyze subtle chemical patterns is unparalleled, making it an indispensable tool in this field [3]. Simply, machine learning is the process by which computers learn patterns in data and apply these patterns to new and unseen data.

## Objective

Predicting the thermomechanical properties of thermoset shape memory polymers (TSMPs) is crucial for their effective design and application. However, the complex topological nature of TSMPs presents a significant challenge in modeling these properties, often resulting in time-consuming processes. To address this challenge and expedite the prediction of the glassy modulus of shape memory polymers (SMPs), we propose the application of machine learning (ML) techniques, aiming to overcome the limitations inherent in traditional modeling methods.

# Methods

In this study, we utilized a small dataset obtained from Ninety percent of the data was used for training and the remainder for validation. various references to train our ML model. Four different datasets were built for each of the models we wanted to predict. In each dataset, four major items were looked out for in the references, namely; the structures of the monomers that are present in the polymers, the weight ratio of each of the monomer and the properties of interest of the respective dataset.

The structure of the monomers was then transformed to the Simplified Molecular Linear Input Line Entry System (SMILES) using an online tool. Smiles denotes a denotes a molecular structure as a graph. The SMILES of the molecules was then fingerprinted. Molecular fingerprinting is a technique used to represent molecules as mathematical objects. This makes it easy to perform statistical studies and/or machine learning methods on the collection of molecules to discover novel insights [4]. Specifically, Morgan fingerprinting was used in this work. This process was done to translate the SMILES input to bit sizes that could be fed as input to the machine learning models. As seen in Fig.

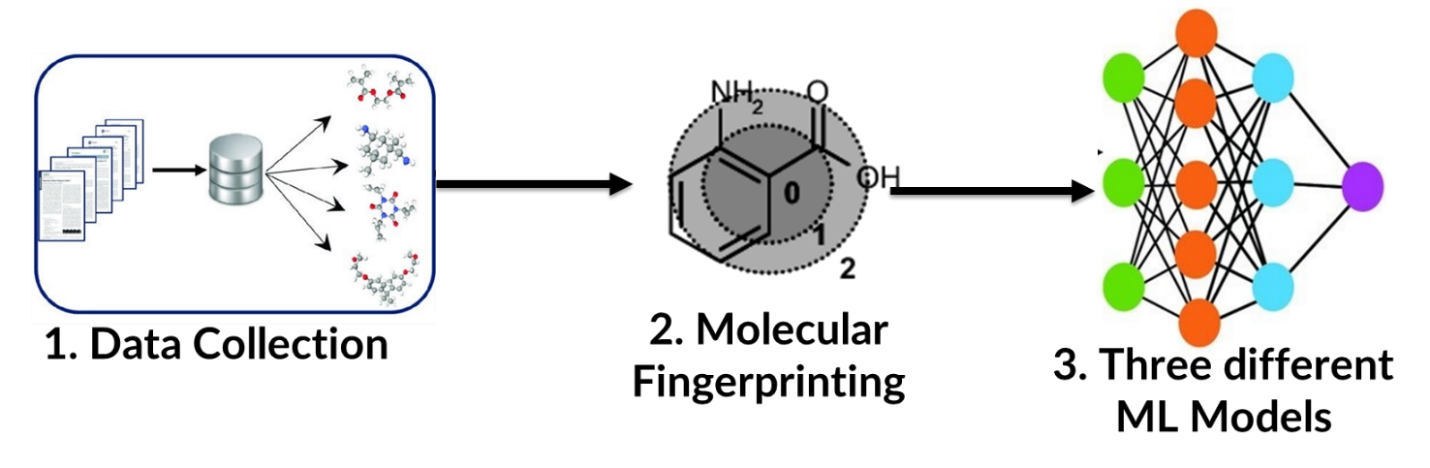


Fig.

The fingerprinted SMILES was the feature and the property for each model was used as the target. For example, in the first model, the glassy modulus the expected output, i.e. the target. Four similar models were built for the four different properties we intended to predict. It must be noted that from here on, each model was named after the respective properties, the model for the glassy modulus will be called glassy model, the one for the rubbery modulus rubbery, model etc.

In order to make a more conclusive decision, three machine learning regression models were compared. The Artificial Neural Network (ANN), Support Vector Regression (SVR) and Gaussian Process Regressors were. These models were selected because …. In each model, the metric of focus was the Mean Absolute Percentage error (see equation 1). The Mean Absolute Error (see equation 2) and Mean Squared error (see equation 3) were used to affirm the case of the best model.

MAPE:

MAE:

MSE:

Where:

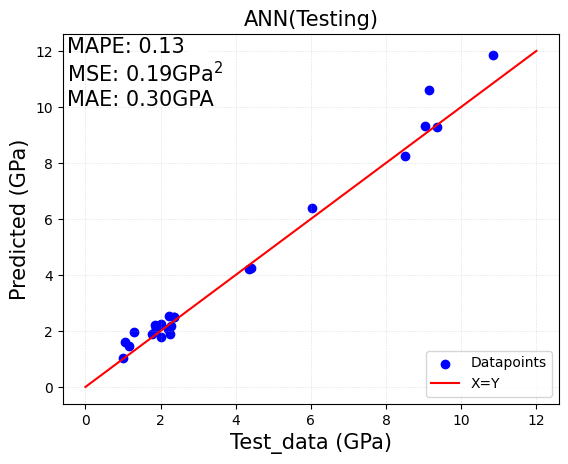
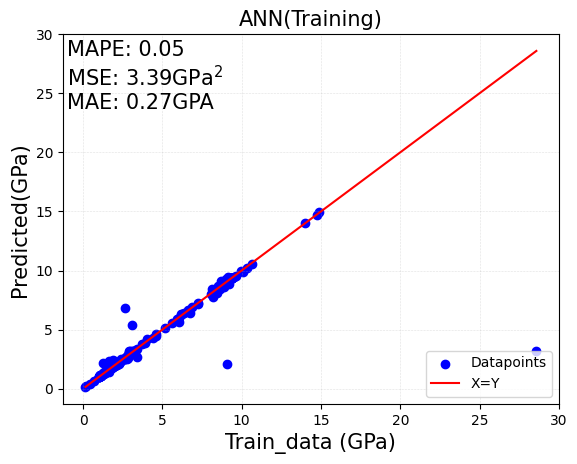
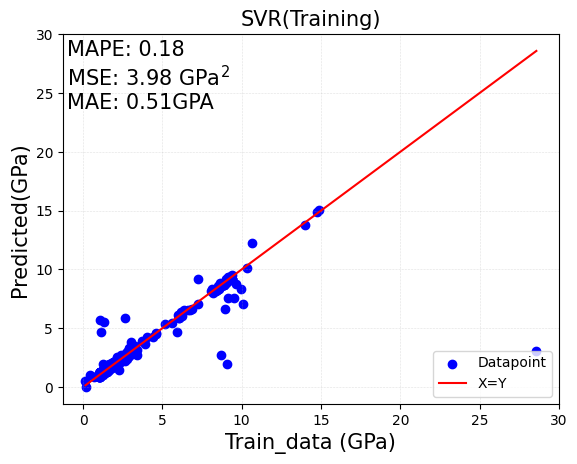
n = number of datapoints

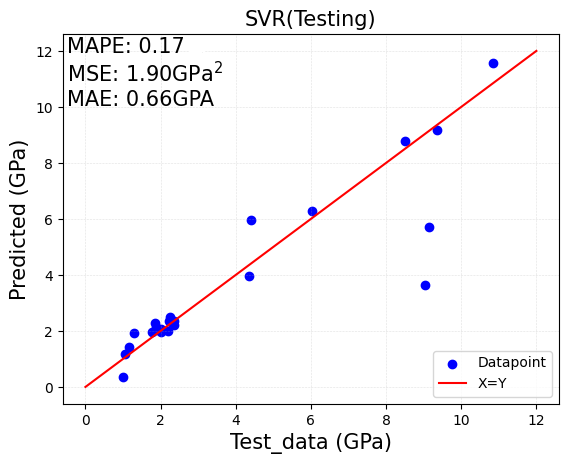
= actual target value at an instance

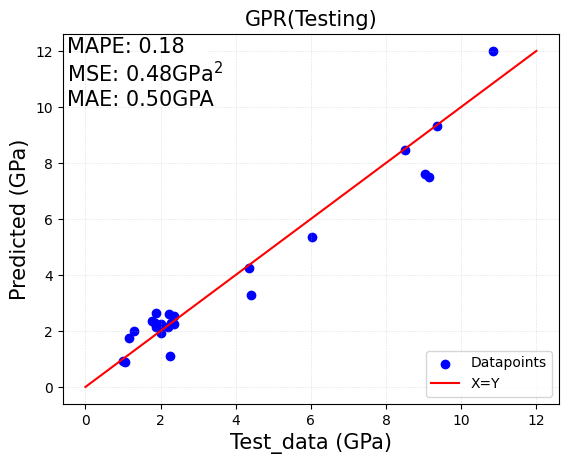
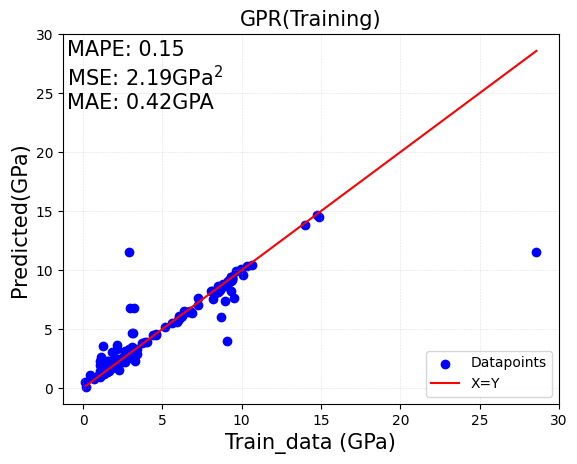
= the predicted target values at an instance

For each model, ninety percent of the data was used for training and the rest for testing. The model parameters were set the same for the four different models to ensure some form of uniformity within the models. Although this was the case, four of the models were not performing as there were expected to. To further ensure that the models were working in an almost similar fashion, the batches were normalized. The L2 regularization was also employed to deal with the issue of overfitting that was occurring.

# Results and Discussions







# Conclusion

# References

[1] S. Gopinath, N. N. Adarsh, P. R. Nair, and S. Mathew, “Shape-Memory Polymer Nanocomposites of Poly(ϵ-caprolactone) with the Polystyrene- block-polybutadiene- block-polystyrene-tri- block Copolymer Encapsulated with Metal Oxides,” *ACS Omega*, vol. 6, no. 9, pp. 6261–6273, Mar. 2021, doi: 10.1021/acsomega.0c05839.

[2] C. Yan, X. Feng, and G. Li, “From Drug Molecules to Thermoset Shape Memory Polymers: A Machine Learning Approach,” *ACS Appl. Mater. Interfaces*, vol. 13, no. 50, pp. 60508–60521, 2021, doi: 10.1021/acsami.1c20947.

[3] M. Aldeghi and C. W. Coley, “A graph representation of molecular ensembles for polymer property prediction,” *Chem. Sci.*, pp. 10486–10498, 2022, doi: 10.1039/d2sc02839e.

[4] Laksh, “A Practical Introduction to the Use of Molecular Fingerprints in Drug Discovery,” 2019. https://towardsdatascience.com/a-practical-introduction-to-the-use-of-molecular-fingerprints-in-drug-discovery-7f15021be2b1 (accessed Mar. 03, 2023).