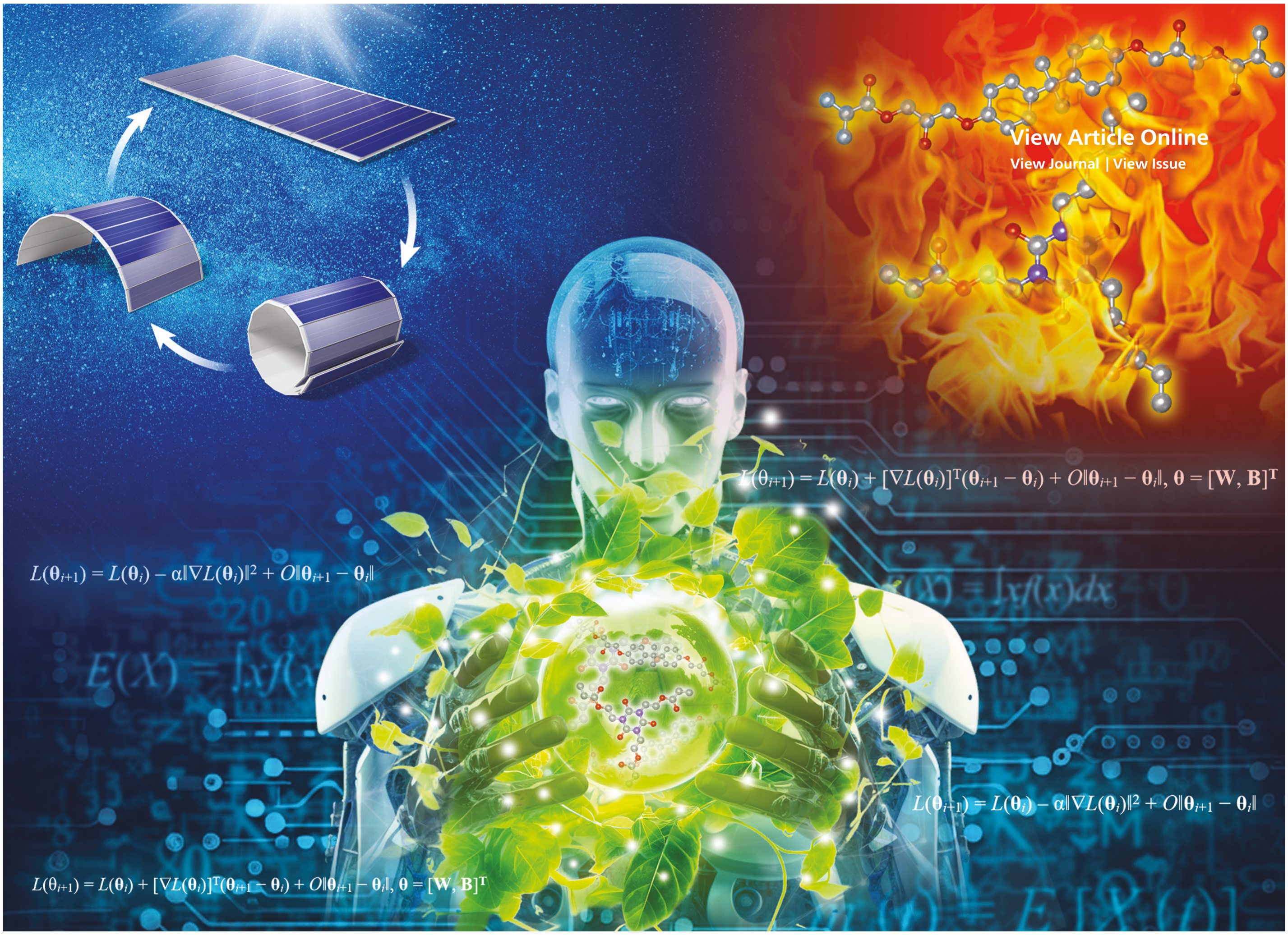
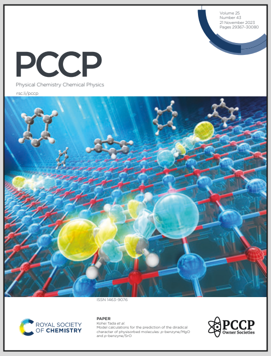


Showcasing research from Professor Cheng Yan and Professor Guoqiang Li's laboratory, Department of Mechanical Engineering, Southern University and A&M College, Baton Rouge, US.

Overcoming the barrier: designing novel thermally robust shape memory vitrimers by establishing a new machine learning framework

This research innovatively applies machine learning (ML) to develop advanced Shape Memory Vitrimers (SMVs). It delves into the ML's mathematical framework, balancing high healing efficiency with elevated glass transition temperatures. Our breakthrough is four thermally robust SMVs, demonstrating exceptional recycling efficiency and shape memory capabilities. A highlight is an SMV with a 233.5 °C Tg, 84.1% recycling efficiency, and 33 MPa recovery stress, showcasing ML's transformative impact in smart polymer design.

As featured in:



See Cheng Yan *et al.*,
Phys. Chem. Chem. Phys.,
2023, **25**, 30049–30065

PCCP

PAPER

Cite this: *Phys. Chem. Chem. Phys.*, 2023, **25**, 30049

Overcoming the barrier: designing novel thermally robust shape memory vitrimers by establishing a new machine learning framework†

Cheng Yan, *[a] Xiaming Feng,[bc] John Konlan,[b] Patrick Mensah[a] and Guoqiang Li [b]

Received 30th July 2023,
Accepted 19th October 2023

DOI: 10.1039/d3cp03631f

rsc.li/pccp

Shape memory vitrimers (SMVs) are an emerging class of advanced materials that have garnered significant interest from researchers in the past five to six years. These materials can return to their original shape when exposed to a stimulus, while also healing damage they have sustained. However, achieving both high healing/recycling efficiency and a high glass transition temperature (T_g) in SMVs has been challenging, due to the conflicting requirements between molecular chain mobility and the formation and reaction of dynamic covalent bond exchange. Based on the understanding of chemo-physical properties, this study first leverages machine learning (ML), involving supervised and unsupervised learning approaches, to navigate this complex design space of SMVs. Furthermore, we elaborated the basic mathematical frameworks of ML approaches and comprehensively compared their performances. Based on the best performing model, we designed four types of thermally robust shape memory vitrimers (TRSMVs), which boast high recycling efficiency, elevated T_g, and exemplary shape memory effects, overcoming conventional barriers. One of the discovered samples exhibited outstanding performance with a T_g of 233.5 °C, a recycling efficiency of 84.1%, and a recovery stress of 33 MPa in experiments. It aligns well with ML predictions, showcasing the potential of our ML framework in driving innovative materials design and advancing the field of smart polymers.

1. Introduction

Based on the response to heat, the formation processes of polymers are different, resulting in two types: thermoplastic polymers and thermosetting polymers. Each type has its own advantages and disadvantages. Thermoplastic polymers have physical crosslinks or entanglement of polymer chains, which allow for easy recycling. However, these polymers are not suitable for applications requiring rubbery solids or high temperature stability, as the physical crosslinks or entangled chains can break down under high temperature conditions. In contrast, thermosetting polymers have stable chemical crosslinks formed by covalent bonds, which grant them excellent rubbery stiffness and thermal stability. Nevertheless, due to the excellent stability of covalent crosslinks, it is difficult to reuse thermosetting polymers. Fortunately, the recent discovery of vitrimers overcomes the limitations of both polymers. According to Zee and Nicolaÿ,[1] vitrimers can empower permanent chemical networks with recycling capability through dynamic covalent bonds, allowing the polymer network to change topology but maintain the number of chemical bonds at temperature below decomposition. Since Montarnal *et al.*[2] found the first vitrimer in 2011, this polymer has gained significant interest and attention in a short time span of about 10 years.[1,3–6] Compared to traditional thermoplastic polymers, vitrimers exhibit higher mechanical strength and longer service life, leading to a reduction in resource waste.[7,8] So far, vitrimers have found wide applications, including shape memory polymers (SMPs),[9,10] 3D printing,[11] in-space assembly for the complex structure,[12] flame-resistant material,[13–15] *etc.*

Synthesis of vitrimers with high recovery efficiency and comparatively low glass transition temperatures (0 °C < T_g < 150 °C) has been widely reported in previous research work.[16–19] However, it remains challenging to increase the T_g of vitrimers while maintaining high recovery efficiency. In previous studies, only a few samples have achieved this goal,[19,20] as demonstrated in Fig. 1. The difficulty in simultaneously achieving high T_g and high recovery efficiency arises

---

[a] *Department of Mechanical Engineering, Southern University and A&M College, Baton Rouge, LA, 70813, USA. E-mail: cheng.yan@sus.edu*

[b] *Department of Mechanical & Industrial Engineering, Louisiana State University, Baton Rouge, LA 70803, USA*

[c] *College of Materials Science and Engineering, Chongqing University, 174 Shazhengjie, Shapingba, Chongqing 400044, China*

† Electronic supplementary information (ESI) available. See DOI: **https://doi.org/10.1039/d3cp03631f**

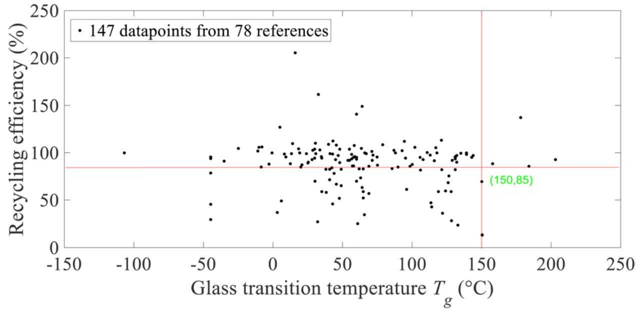


Fig. 1 T_g *vs.* recycling efficiency. It is evident that only a few samples can exhibit both high T_g (>150 °C) and excellent recovery efficiency (>85%).

from the fact that a high T_g needs low mobility of polymeric chains and segments, while high healing/recovering efficiency needs high mobility of the polymeric chains and segments. In other words, these high T_g polymers exhibit low mobility, which is unfavorable for the formation and reaction of dynamic covalent bond exchange. Another challenge is that the intrinsic healing of vitrimers depends on the formation of reversible dynamic covalent bonds at a certain high temperature, which is the topology freezing transition temperature (T_v). Usually, T_v is higher than T_g. Therefore, vitrimers with high T_g may make the T_v close to or even higher than the decomposition temperature, leading to difficulty in maintaining self-healing. Nevertheless, developing vitrimers with high T_g and high recovery efficiency would significantly broaden their application windows. For example, vitrimer-made structures are able to perform self-healing in a wider temperature range and perform under heavy load-carrying conditions, as high T_g vitrimers typically possess high strength. Additionally, polymers with high T_g offer notable advantages in outer space. For example, the temperature on the surface of Mars can reach 120 °C, while on Venus, the temperature can even soar up to 462 °C. Polymers with high T_g can maintain their mechanical strength under these extreme conditions, making them highly desirable for potential future colonization on other planets and presenting a promising application outlook. Meanwhile, tough working conditions in outer space often necessitate polymers to exhibit multiple functionalities. As such, we aim to design vitrimers with exceptional shape recovery performance, allowing them to assume two different shapes upon actuation. This shape memory effect is desired especially in their application in deployable structures[21] or in damage self-healing.[22] In addition to space structures, shape memory polymers with high T_g are highly desired as proppants or lost circulation materials in the oil and gas and geothermal drilling applications.[23,24] In this study, we refer to these new materials as "thermally robust shape memory vitrimers (TRSMVs)".

Due to the conflicting design requirement between the mobility of molecular segments and glass transition, it is challenging for the design of TRSMVs. Although innovative materials such as shape memory polymers,[25] vitrimers,[2] piezoelectric materials[26] have been developed, creating cutting-edge materials typically requires significant talent, time, domain knowledge, and a bit of luck, rendering the process difficult to perform and slow. Particularly, when we design TRSMVs, the different properties and functionalities, such as high thermal stability, shape memory, and dynamic covalent bonds, may reside in different composition spaces. Identifying the overlapping compositional space becomes more challenging. Hence, developing a comparatively easy material design method is highly desired.

According to Yan and Li,[27] the key for polymer design is to understand the relationship between microscopical structures and macroscopical properties, such that

$$S = \phi(\mathbf{w}),\ \mathbf{w} = \arg\min_{\mathbf{w}} L(\mathbf{w}) \tag{1}$$

where S represents the appropriate polymer structure, $\mathbf{w}$ is the input tensor, and L is the loss function or objective function.

Although the structure–property relationship is well-defined since all the mechanical behaviors follow Newton's laws of motion, quantifying the molecular and topological structures of polymers is challenging, and the computational cost is often significantly high.[28] Specifically, different from general properties, both the experiments for dynamic bond exchange and shape memory effect need a couple of hours to one day, while the molecular dynamic (MD) approach can only accurately simulate atomic-level mechanisms of vitrimers on a nanosecond scale (limited by current computational capacity). This disparity in timescales results in MD simulations being considerably misaligned with experimental results. For instance, there is a significant discrepancy of 223% between the MD prediction of recovery stress, which is 55 MPa,[29] and the experimental result of 17 MPa.[30] Hence, employing ML approaches to model the structure–property relationship has become a wise and attractive solution.

The recent rapid development in both hardware and algorithm for ML has opened up new possibilities for material discovery. Over the last decade, there has been a surge in the discovery of new materials, thanks to the application of ML across various domains.[31–33] Specifically, in the realm of SMPs, addressing the challenge of limited datasets, Yan and his collaborators innovatively employed a dual-convolutional network[34] and a transfer learning-variational autoencoder (TL-VAE)[35] in their research. This led to the discovery of two groundbreaking thermoset shape memory polymers (TSMPs), both surpassing existing performance metrics for SMPs. In a subsequent cutting-edge development, Yan and colleagues introduced a novel self-enhanced deep neural network (SDNN) by first incorporating a self-attention algorithm for the prediction of flame retardancy, resulting in a highly efficacious ML model.[36] In the field of metals and alloys, Lee *et al.* designed a MVAE framework that is able to generate 10 new alloy entries with higher yield strength and ultimate tensile strength.[37] In the field of metamaterials, Adithya *et al.* designed new cellular structures and lattice structures with improved properties (load-carrying capacity, natural frequency, energy absorption capabilities) using a generative adversarial network (GAN).[38,39] Despite the previous successes, no ML-based methods have been applied to the field of vitrimer discovery, hence we decided to employ a design framework based on ML to address the gap in this field.

In our study, three new innovative strategies are employed to design new TRSMVs, which are our important contributions. First, by comparing the properties and performance, we selected the best performing model from four ML models, which was based on our understanding of different properties. Second, in the absence of any open database for both vitrimers and SMPs, we manually gathered data points from references and created three different databases. Lastly, to perform effective virtual screening for TRSMVs, we developed a set of novel screening criteria for virtual material screening, aiming to overcome the barriers imposed by small datasets.

This article is structured as follows. Section 2 begins with outlining a material discovery framework for TRSMVs. We then proceed with a qualitative analysis to identify the key features for fingerprinting TRSMVs, and further discussed these fingerprinting methods and ML prediction models. After that, we leveraged four different ML approaches to predict the properties of TRSMVs, comparing their respective performances. Furthermore, based on the best-performing ML model, we identified several promising new TRSMVs and experimentally validated one of them. Section 3 provides important conclusions drawn from our study.

2. Method

2.1 Machine learning framework

The framework to discover TRSMVs is shown in Fig. 2. It is primarily composed of two parts, *i.e.*, forward prediction and inverse mining. Beginning with forward prediction, it operates based on the data gathered from references shown in Fig. 2a(1).

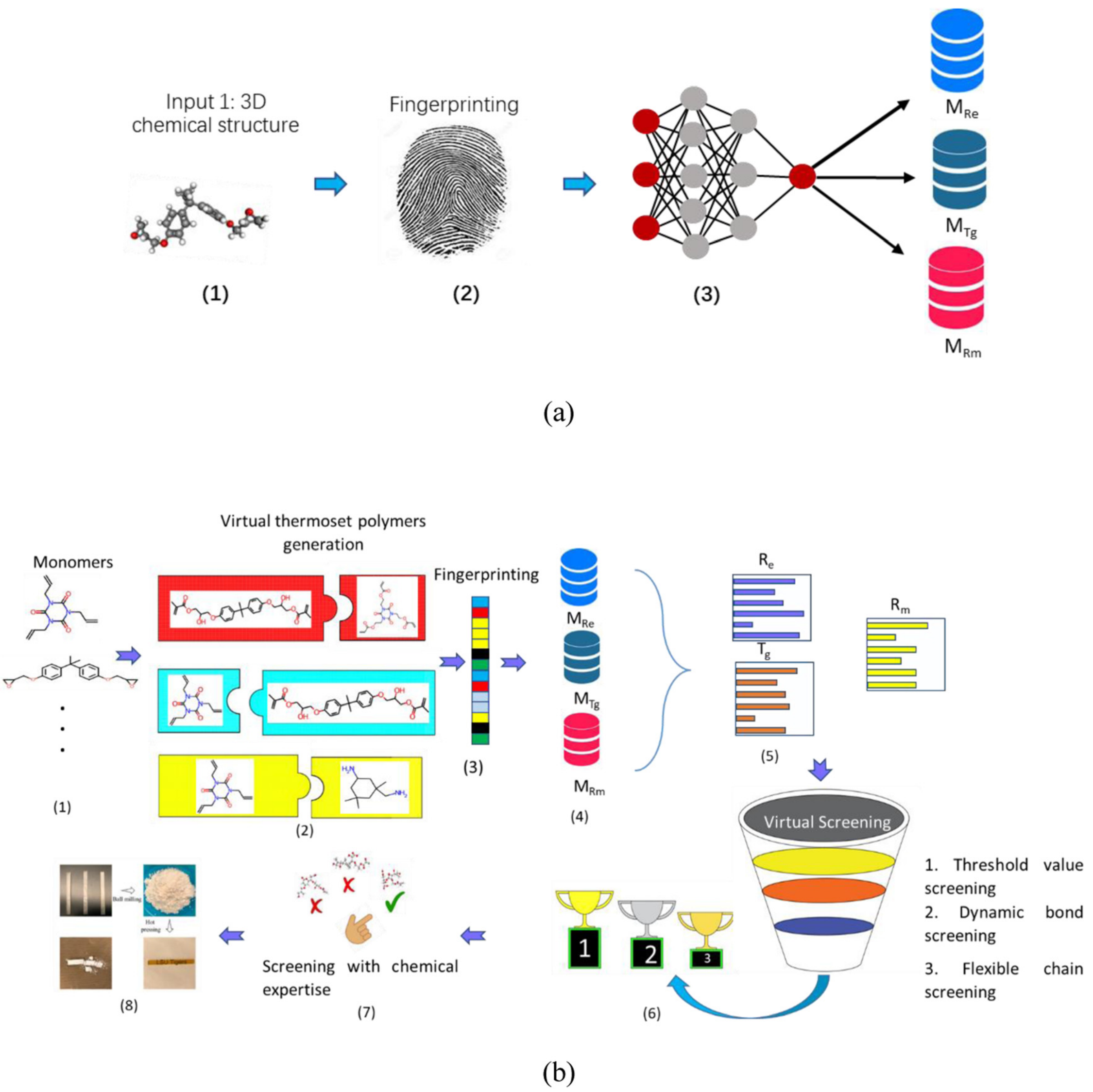


Fig. 2 Framework to discover TRSMVs: (a) forward prediction and (b) inverse mining.

This element of the framework involves training three ML models *i.e.*, M_{R_e} (model for recovery efficiency R_e), M_{T_g} (model for glass transition temperature T_g), M_{R_m} (model for rubbery modulus R_m), as shown in Fig. 2a(3). The training process employs an appropriate fingerprinting method (Fig. 2a(2)), and its details are presented in Section 2.3.

Subsequently, an inverse mining component is employed. Its objective is to search for the desired samples through the established database. This involves eight steps (see Fig. 2(b)), details of which are elucidated as below:

(1) Gather monomers and crosslinkers from references to create databases for SMVs.

(2) Randomly pair monomers and crosslinkers capable of polymerization to produce potential SMPs.

(3) Represent potential SMPs as Simplified Molecular Input Line Entry System (SMILES) strings for fingerprinting purposes.

(4) Input these SMILES groups into pre-trained machine learning models for recycling efficiency, T_g, and rubbery modulus. Evaluate and select the best-performing models comprehensively.

(5) Predict T_g, R_e and R_m of all potential TRSMVs using chosen models.

(6) First, the SMPs with functional groups that can lead to a bond exchange reaction will initially be selected as targets. Next, apply a set of screening criteria to remove retardant targets. Finally, identify material champions for TRSMVs.

(7) Further screen material champions using chemical knowledge.

(8) Validate the TRSMVs through a series of experimental testing.

2.2 Key features analysis

In our model design, we aim to predict three targets in the model, including T_g, recycling efficiency, and recovery stress. These three targets involve a combination of intricate chemical and physical processes, necessitating a thorough understanding to select the suitable fingerprinting method.

T_g is primarily influenced by factors such as the polymer structure, molecular weight, intermolecular forces, crosslink density, and degree of polymerization. Among them, the most pivotal factor is segment mobility, which is determined by the polymer structure. For example, rigid chains (*e.g.*, benzene ring) result in high T_g, while soft chains (*e.g.*, long aliphatic chains) lead to low T_g. Nonetheless, the current discussion provides only an approximated, qualitative description of T_g, selectively omitting several factors for the sake of brevity. As presented above, the molecular mass has an important role in determining T_g, a relationship that has been rigorously investigated and validated by prior studies. For instance, the works of Novikov and Rössler,[40] as well as Wolf *et al.*,[41] have elucidated the quantitative relationship between T_g and molecular mass. Meanwhile, Zhang and Xu[42] have employed molecular traceless quadrupole moment and molecule average hexadecapole moment in a machine learning model to predict T_g, marking another notable approach in this multifaceted exploration. Furthermore, it's imperative to acknowledge that T_g is influenced not only by the static factors mentioned but is also significantly affected by thermodynamic state changes, a concept explored by the Random First-Order Transition (RFOT) theory.[43] Recycling efficiency hinges on three decisive factors. First, the monomers used for vitrimer synthesis play a vital role, as they directly determine the number of dynamic bonds. Polymer networks with more dynamic bonds generally yield more dynamic chemical reactions, thus improving recycling efficiency. Taking phosphate vitrimer as an example, a phosphate vitrimer synthesized using phosphoric acid 2-hydroxyethyl methacrylate ester and EPON-826[44] has higher recycling efficiency than the vitrimer synthesized by bis[2-(methacryloyloxy) ethyl]phosphate[15] due to a larger number of dynamic bonds in phosphates (99% *vs.* 75.5%). Specifically, phosphoric acid 2-hydroxyethyl methacrylate ester has three dynamic bonds, while bis[2-(methacryloyloxy)ethyl]phosphate has one dynamic bond, hence the former can lead to more recycling efficiency under suitable reaction conditions. Secondly, recycling efficiency is partially determined by T_g; a low T_g requires a lower recycling temperature, resulting in less aging, hence leading to high recycling efficiency. Specifically, the recycling temperature of vitrimers should exceed their T_g, establishing that a lower T_g typically corresponds to a lower recycling temperature. In this context, operating at a diminished recycling temperature decreases the likelihood of chain scission and cross-linking, thereby ensuring a minimal attenuation of the mechanical properties of SMVs. Concurrently, reduced aging sustains superior molecular mobility and potentially lessens the packing density in the polymeric glass, thereby amplifying reactivity during the chemical recycling process. For example, epoxy vitrimers were synthesized by curing bisphenol A diglycidyl ether (DGEBA) with adipic acid (AA) possessing a low T_g = 47°, hence resulting in a high recycling efficiency R_e = 97%.[45] In contrast, a SMV formulated using bisphenol A glycerolate dimethacrylate (BPAGMA) exhibits a T_g of 150°, which corresponds to a lower recycling efficiency R_e = 69.5%.[10] Lastly, activation energy plays a role in determining recycling efficiency. A lower activation energy means a less energy threshold to activate a dynamic bond, rendering more occurrences of dynamic reactions and improving the recycling efficiency. Activation energy is typically determined by the types of functional groups and network structures. For instance, Zhu *et al.* synthesized a vinylogous urethane vitrimer with acetoacetylated castor oil (ACO) and aminated DL-limonene (ADL). This vitrimer has a low activation energy of 26 kJ mol^{-1}, resulting in a relatively high recycling efficiency E_{re} of 103.63%.[46] On the other hand, Ma *et al.* developed a vitrimer named MDS-EPO using 4,40-methylenedianiline (MDA) and IS-EPO.[47] This vitrimer possesses an activation energy of 106 kJ mol^{-1} and displays a medium recycling efficiency E_{re} of 82.6%.

Recovery stress prediction is challenging due to the interplay between experimental conditions and microscopic structures. However, based on general experience, a higher rubbery modulus correlates with higher recovery stress. Hence, we employed a rubbery modulus prediction model as a surrogate for the recovery stress prediction model, a method proven effective in previous works.[35] The rubbery modulus is primarily

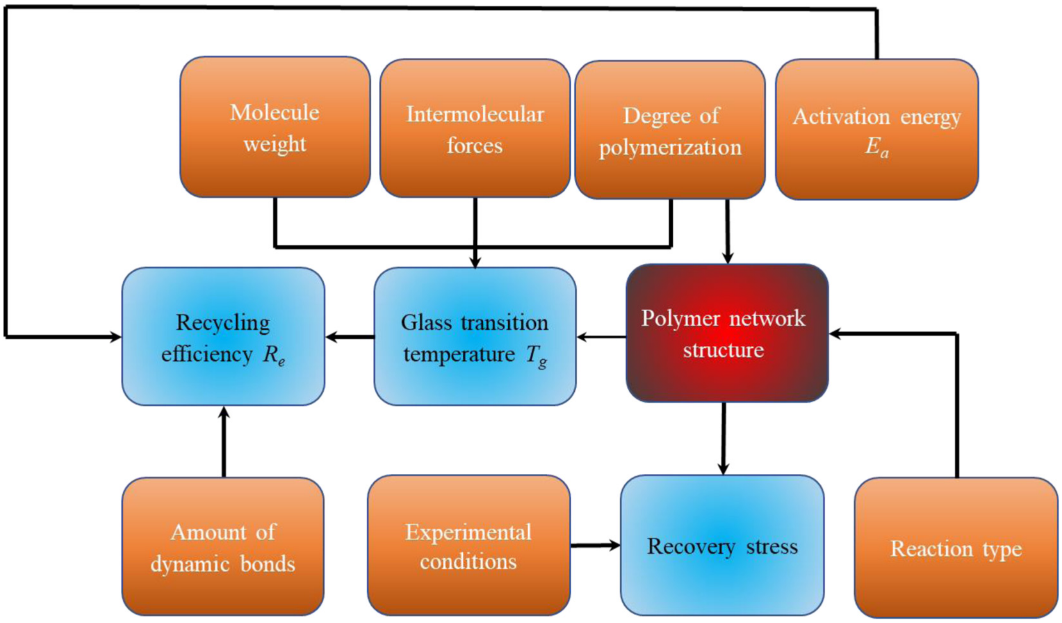


Fig. 3 Correlation between basic features and ground truths.

determined by the number of rigid chains in the polymer network, and this is also viewed as macroscopic expression arising from physical behavior of the underlying macroscopic structures. Fig. 3 displays the qualitative correlations for all the three prediction targets. Upon reviewing the three prediction targets, it becomes apparent that the molecular structure is the most crucial factor for all of them, followed by functional group rank. As a result, our feature extraction should prioritize by capturing these structural details. In ML, the Simplified Molecular Input Line Entry System (SMILES) is a widely used chemical linear notation for representing 3D molecular structures. While the SMILES effectively captures the primary elements and bond types, it has limitations in representing complex structures. Direct SMILES encoding places great emphasis on individual elements and is highly dependent on underlying dictionary.[35] As a result, it may not yield satisfactory performance for model prediction. Consequently, it is essential to select a fingerprinting method that is able to further capture more topological structures of a polymer network. To account for functional group factors, we employed multiple screening approaches to enhance our model prediction, which will be demonstrated in Section 2.5.2.

2.3 Fingerprinting methods

In this study, we applied two different feature extraction approaches: (1) Transferred Variational Autoencoder (TL-VAE) and (2) Morgan fingerprinting. Detailed descriptions of each approach are provided below.

2.3.1 TL-VAE model. In a previous study, Yan *et al.*[35] demonstrated that the Transferred TL-VAE is a better option than the direct SMILES encoding for featurization of polymer network structures. In this study, we continue to employ this feature extraction method. The TL-VAE is a reasonable choice for two reasons. First, the input of the TL-VAE is derived from the SMILES, making it naturally capture the features of primary elements and functional groups. When coupled with the molar ratio between monomers and crosslinkers used in polymer synthesis (detailed in Section 2.5.1), this method can effectively capture the type of functional groups and reaction types. Second, the VAE often requires a large amount of data points. In our previous study, it was shown that drug molecules have exhibited similar structures with monomers or crosslinkers of SMPs. Therefore, in our case, the TL-VAE model can also resort to this large dataset of drug molecules, rendering it an appropriate approach for our purpose. It should be mentioned that we have adopted two parameter optimizations for the TL-VAE model, including the dimension of latent space and loss function (see supporting information in our previous publication[35]). To implement this model in our current study, we first trained the model with 420 000 drug molecules and then trained the last two layers with 389 monomers or crosslinkers from SMPs or vitrimers (see Fig. 4). The rationale for employing Transfer Learning Variational Autoencoders (TL-VAEs) can be appreciated from three perspectives. First, the TL-VAE must effectively grasp the syntax of the SMILES, a linear notation system that can be challenging, particularly when dealing with monomers featuring multiple ring structures. For instance, accurately mapping the high-dimensional vector of 4′-diaminodiphenyl sulfone (DDS), which has four rings, back to its original SMILES representation is a complex task. The SMILES notation for DDS is "C=CCc4ccc(OCC1CO1)c(c2cc(CC=C)ccc2OCC3CO3)c4". To achieve precise mapping, the VAE model must correctly identify the number of rings and match all closing brackets, a task that often poses a challenge to this model. Second, the crosslinkers and monomers of SMVs possess unique functional groups not found in drug molecules, necessitating fine-tuning of the

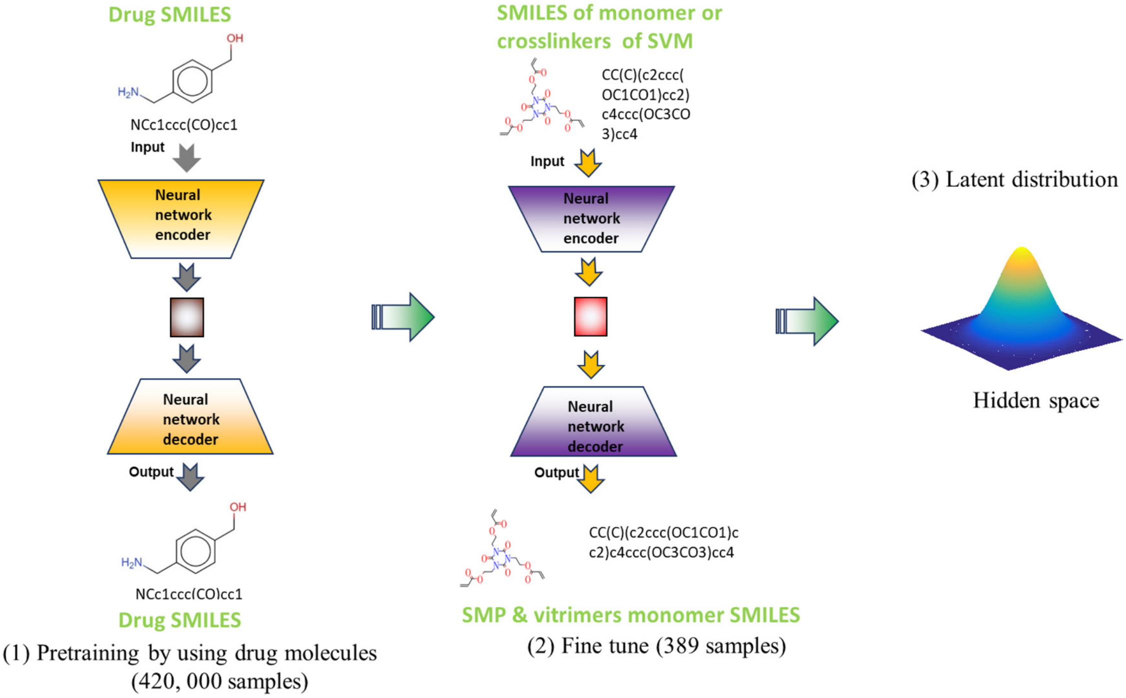


Fig. 4 Training strategy for the network of the VAE model. The detailed model pipeline can be found in the ESI.†

model. Third, VAEs typically involve a large number of fitting parameters (our model includes 1 087 565 parameters), making it essential to train it with sufficient data points. Otherwise, we inevitably encounter overfitting. As illustrated in Fig. 4, the 389 SMP monomer or vitrimer monomer SMILES is insufficient for our TL-VAE model.

VAEs can be mainly divided into encoders and decoders. The encoder maps inputs into vectors in a hidden space, which can be written as

$$E = q(z|x) \tag{2}$$

where x and z represent the input and the vector in the hidden space. In contrast, the decoder maps the vector in the hidden space back to the inputs, such that

$$D = q(x|z) \tag{3}$$

The aim of the VAE model is to render the output close to the input as much as possible, hence we leverage "categorical cross entropy" as the loss function, which can be written as

$$L(p,q) = -\sum_{1}^{N} p(x)\log q(x) \tag{4}$$

where $p(x)$ is the expected value and $q(x)$ is the predicted probability for the expected value, and N is the number of categories, *i.e.*, the number of symbols in the dictionary for the SMILES in a batch.

The VAE model is composed of four types of layers, *i.e.*, convolutional layer (CNN), long-short term memory layer (LSTM), flatten layer and fully connected layer (FL). Among them, the CNN layer aims at extracting features from the binary image representing SMILES, which can be written as

$$\mathbf{g} = \mathbf{f} * \mathbf{h} \tag{5}$$

where $\mathbf{f}$ is the filter tensor, $\mathbf{h}$ is the image tensor, and "*" represents the convolution operation. The piecewise function "Relu" and piecewise function "Softmax" are the activation functions used in the network and which read as

$$\sigma_{\mathrm{Re}}(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases} \tag{6}$$

$$\sigma_{\mathrm{s}}(\mathbf{x})_i = \frac{\mathrm{e}^{x_i}}{\sum_{j=1}^{k} \mathrm{e}^{x_j}} \tag{7}$$

A LSTM is a type of recurrent neural network (RNN). The basic cell for a LSTM is shown in Fig. 5. It is believed that the current state is dependent not only on the last state but also on the previous history, so it defines two types of hidden states, namely c_t and h_t. With input vector $\mathbf{x}_t$ in ith state and the last hidden state inputs c_{t-1} and h_{t-1}, the intermediate vector $\mathbf{Q}_t$ (which can be $\boldsymbol{f}_t$, $\boldsymbol{i}_t$, $\boldsymbol{g}_t$, $\boldsymbol{o}_t$, see Fig. 5) can be calculated using the same method as

$$\mathbf{Q}_t = \phi(\mathbf{W}_t\mathbf{x}_t + \mathbf{U}_t\mathbf{h}_{t-1} + \mathbf{B}_t) \tag{8}$$

where $\mathbf{I}$ is the identity tensor, $\mathbf{W}_t$, $\mathbf{U}_t$ and $\mathbf{B}_t$ are the weight tensor for an input, the weight tensor for hidden state and the bias vector, respectively. Evidently, a total of four weight tensors and four bias tensors are calculated. It is important to note that the calculation between weights and inputs involves matrix multiplication. ϕ is an

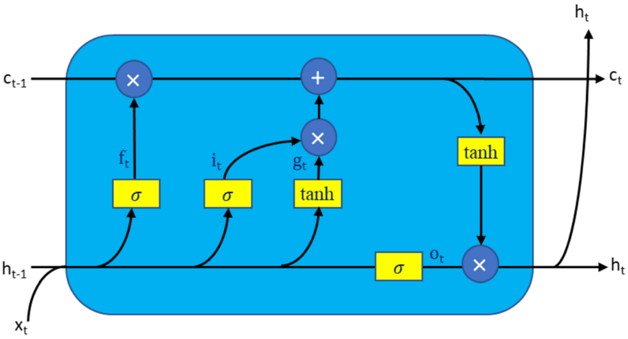


Fig. 5 The basic structure of a LSTM cell.

activation function, which can be either the logistic function (σ) or the hyperbolic function (tanh) and is read as

$$\sigma(x) = \frac{1}{1+\mathrm{e}^{-x}} \tag{9}$$

$$\tanh(x) = \frac{\mathrm{e}^{2x}-1}{\mathrm{e}^{2x}+1} \tag{10}$$

The two hidden states c_t and h_t can be calculated using

$$c_t = \mathbf{f}_t \odot \mathbf{g}_{t-1} + \mathbf{i}_t \odot \mathbf{g}_t,\ h_t = \mathbf{o}_t \odot \tanh(\mathbf{g}_t) \tag{11}$$

where $\odot$ represents the Hadamard product.

Based on the network architecture we have designed (refer to the ESI† for details), the loss function (as per eqn (4)) can be characterized as a function with a multitude of parameters. Our objective is to accurately estimate these parameters. To achieve this, we employ Taylor's theorem. Consequently, the loss function or objective function at the $(i+1)$th step can be expressed as a function of weights and biases[48]

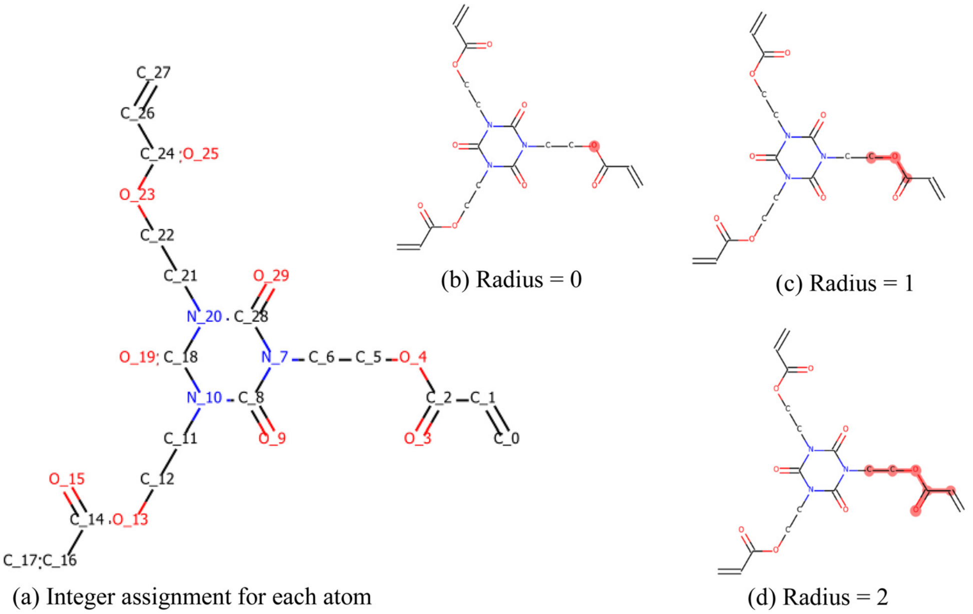
$$L(\theta_{i+1}) = L(\boldsymbol{\theta}_i) + [\nabla L(\boldsymbol{\theta}_i)]^{\mathrm{T}}(\boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i) + O\|\boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i\|,\ \boldsymbol{\theta} = [\mathbf{W},\mathbf{B}]^{\mathrm{T}} \tag{12}$$

where $\boldsymbol{\theta}$ is the updatable tensor (which represents weight tensor $\mathbf{W}$ and bias tensor $\mathbf{b}$). Stipulating a positive learning rate α (a constant) and let $\boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i = -\alpha\nabla L(\boldsymbol{\theta}_i)$, then the loss function eqn (12) can be written as

$$L(\boldsymbol{\theta}_{i+1}) = L(\boldsymbol{\theta}_i) - \alpha\|\nabla L(\boldsymbol{\theta}_i)\|^2 + O\|\boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i\| \tag{13}$$

Apparently, the loss function continues to reduce with iterative updating of θ. In other words, the comprehensive error between prediction and ground truth can be minimized. This model took about 20 hours for training in our workstation equipped with RTX2080 Ti.

2.3.2 Morgan fingerprinting. In addition, we utilized Morgan fingerprinting to extract features, which has been commonly recognized as an effective finger printing approach for structure featurization. We implemented this method for comparison purposes. Morgan fingerprinting is also good at describing polymer network structures. The method was originally proposed by Rogers and Hahn,[49] which can be simply described as follows. First, using the daylight atomic invariants rule,[50] identity information is assigned to each non-hydrogen atom in the molecule, *i.e.*, atomic number, atomic mass, number of nearest-neighbor atoms, *etc.* Secondly, bond information around an atom is described based on the first step, which generates a new vector group for each atom and can then produce a new identity using a hash function. Through multiple iteration (in each iteration, radius expands by 1, as shown in Fig. 6), every atom and their neighbors are finally fingerprinted as a new vector with desired length. In this process, the repeated molecule fragments will be deleted, and the radius



(a) Integer assignment for each atom

(b) Radius = 0

(c) Radius = 1

(d) Radius = 2

Fig. 6 A typical iteration process for generating the Morgan fingerprinting tris[2-(acryloyloxy)ethyl]isocyanurate, a monomer molecule of the shape memory polymer. Starting with element 4 (oxygen) as the center, the Morgan fingerprinting progressively expands its radius to represent molecular structures (see the highlighted regions). With $R = 13$, the iteration process will count all the elements.

range should be defined according to the complexity extent of the molecules. Considering that some monomers possess a couple of benzene rings, we employed a larger radius, *i.e.*, $R = 13$, which aims to avoid the potential confusions due to a smaller radius definition.

2.3.3 TSMP fingerprinting. By utilizing the Variational Autoencoder (VAE) model or Morgan fingerprinting, it is possible to extract features from any monomer or crosslinker through the use of an encoder. Subsequently, a thermoset shape memory polymer (TSMP) network can be characterized through fingerprinting. Let's consider that a TSMP network is synthesized using 'n' distinct types of monomers or crosslinkers. These can be represented as vectors, denoted as $m_1, m_2, \ldots, m_n$. In parallel, their corresponding molar percentages are represented as $a_1, a_2, \ldots, a_n$. Consequently, this polymer network can be depicted as a high-dimensional resultant vector as

$$\mathbf{S} = \mathbf{m}_1 \cdot a_1 + \mathbf{m}_2 \cdot a_2 + \mathbf{m}_3 \cdot a_3 + \ldots \mathbf{m}_n \cdot a_n \tag{14}$$

This vector will be used in the following forward prediction model as features for input.

2.4 ML model

For the forward prediction model, we leveraged two types of ML models in this study, *i.e.*, artificial neural network (ANN) and support vector method (SVM).[51] Their basic math frameworks are illustrated as follows.

2.4.1 ANN model. ANN is a type of engineering optimization method and aims to solve a minimization problem through the high-dimensional Taylor theorem. The network architecture is shown in Fig. 7. In a regression problem, we employed the activation functions "ReLU" and "linear" function, which can be written as

$$\mathrm{ReLU}(x) = \begin{cases} x & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases} \quad \mathrm{Linear}(x) = x \tag{15}$$

The function for output can be represented as

$$f(\mathbf{X}) = \mathrm{Linear}(\mathbf{W_7} \cdot \mathrm{ReLU}(\mathbf{W_6} \cdot \mathrm{ReLU}(\mathbf{W_5} \cdot \mathrm{ReLU} \times (\mathbf{W_4} \cdot \mathrm{ReLU}(\mathbf{W_3} \mathrm{ReLU}(\mathbf{W_2} \times \mathrm{ReLU}(\mathbf{W_1 X} + \mathbf{b_1}) + \mathbf{b_2}) + \mathbf{b_3}) + \mathbf{b_4}) + \mathbf{b_5}) + \mathbf{b_6}) + \mathbf{b_7}) \tag{16}$$

The loss function that we chose is the "mean average percentage error (MPAE)", which can be expressed as

$$L(\theta) = \frac{\sum_{1}^{i} \left[y - \hat{y}_i \right]}{i} \tag{17}$$

where y_i and y_i are prediction outcome and ground truth, respectively. The loss can be reduced using eqn (13). Also, the weights and bias in the ANN model can also be optimized following $\boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i = -\alpha \nabla L(\boldsymbol{\theta}_i)$ (learn rate $\boldsymbol{\alpha}$ can be found in Table 1).

2.4.2 Support vector regression (SVR). Another ML method that we used here is SVR. The structure–property relation for SVR yields a non-linear relation given by[52]

$$f(\mathbf{x}) = \mathbf{w} \cdot \phi(\mathbf{x}) + b \text{ with } \mathbf{w} = \sum_{i=1}^{N} (\alpha_i^{+} - \alpha_i^{-}) \phi(\mathbf{x}_i) \tag{18}$$

where α_i and α_i^* are Lagrange multipliers, and b is bias. $\phi(\mathbf{x})$ is the feature mapping function. The dual form of eqn (18) can be written as

$$f(\mathbf{x}) = \sum_{i=1}^{N} (\alpha_i^{+} - \alpha_i^{-}) K(\mathbf{x}_i, \mathbf{x}) + b \tag{19}$$

The kernel function K is the radial basis function (RBF), which can be expressed as

$$K(\mathbf{x}_i, \mathbf{x}) = \exp(-\|\mathbf{x}_i - \mathbf{x}\|^2) \tag{20}$$

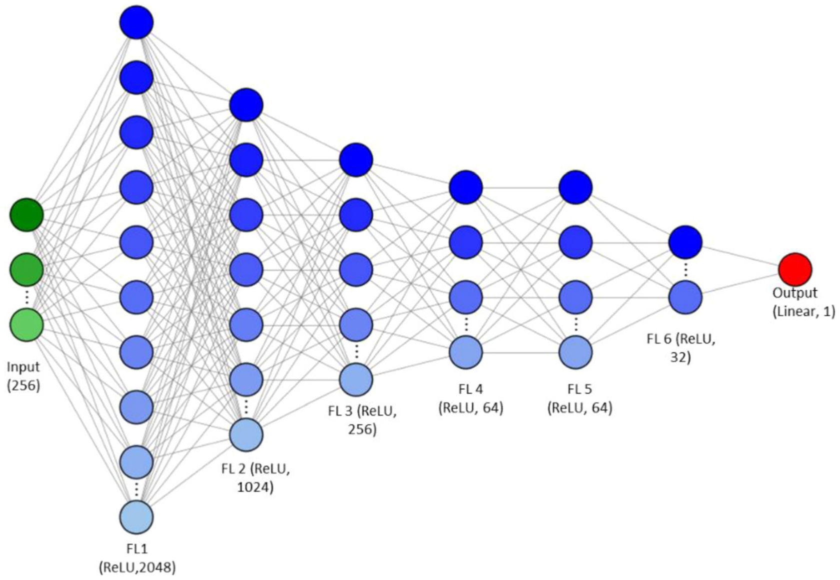


Fig. 7 Basic pipeline structures for the artificial neural network.

Table 1 Hyperparameters adopted in the training of the ANN model

Hyperparameter names	Values or items
Ratio between training data and test data	80/20
Batch size	32
Learning rate	0.01
Activation function for 1–5 hidden layers	ReLU
Activation function for the output layer	Linear
Neuron number in hidden layers 1–6	2048, 1024, 256, 64, 64, 32
Random state in T_g, R_e, R_m	24, 1, 7

The objective function is to minimize the error outside the ε-insensitive tube and model complexity, which can be expressed as a function as shown below,

$$\min\left(C\sum_{i=1}^{N}\left(\xi_i^{+}+\xi_i^{-}\right)+\frac{1}{2}\sum_{i,j}\|\mathbf{w}\|^2\right) \text{ subject to } t_i \le y_i+\varepsilon+\xi^{+} \text{ and } t_i \le y_i-\varepsilon+\xi^{-} \tag{21}$$

where C is a constant, ξ_i^+ and ξ_i^- are the upper boundary and lower boundary, respectively. N is the total number of samples, and ε is a small margin. This objective function can be written as the dual formula by introducing Lagrange multipliers α_i^+, α_i^- as shown below,

$$\max_{\alpha_i^{+},\alpha_i^{-}}\left[\sum_{i=1}^{N}\left(\alpha_i^{+}-\alpha_i^{-}\right)t_i-\varepsilon\sum_{i=1}^{N}\left(\alpha_i^{+}-\alpha_i^{-}\right)-\frac{1}{2}\sum_{i,j}\left(\alpha_i^{+}-\alpha_i^{-}\right)\left(\alpha_j^{+}-\alpha_j^{-}\right)K\left(\mathbf{x}_i,\mathbf{x}_j\right)\right] \tag{22}$$

subject to

$$0 \le a_i^{+} \le C,\ 0 \le a_i^{-} \le C \text{ and } \sum_{i=1}^{N}\left(a_i^{+}-a_i^{-}\right)=0\ \forall_i \tag{23}$$

where N is the total number of samples. Our aim is to solve the Lagrange multipliers and bias. By leveraging quadratic programming, this problem can be quickly solved. This can be realized using a software machine learning library, Scikit-learn.[53]

It is apparent that both of the aforementioned approaches involve intricate iterative computations. However, leveraging the robust parallel processing capabilities of Graphics Processing Units (GPUs), these models (ANN and SVR) can be efficiently executed within a matter of minutes.

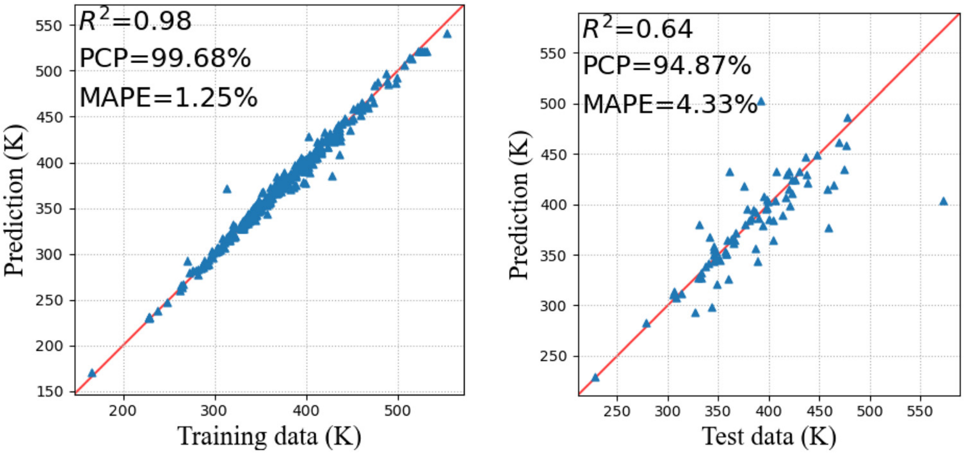
2.4.3 Result comparison. By integrating two fingerprinting methods and two supervised learning methods, we employed four unique approaches to predict T_g, rubbery modulus and recycling efficiency, respectively. The results are exhibited in Table 2 and Fig. 8 and 9, respectively. Herein, we introduced three different indices, including mean average percentage error (MAPE), coefficient of determination (R^2) and percentage of correct point (PCP) to evaluate the performance of the models. Given more datapoints for T_g and less data points for rubbery modulus and recycling efficiency in the training dataset, we established a prediction accuracy threshold. For T_g, prediction with an error of 15% or less was considered accurate, while for rubbery modulus and recycling efficiency, prediction with an error of 20% or less was deemed accurate. Upon thoroughly comparing with the four approaches, we found that the approaches "VAE + ANN" and "Morgan encoding + ANN" exhibit the best performance among the group. We can mathematically understand this outcome by examining the forms of fitting functions and fitting parameters. First, the assumed function for the ANN appears to be better than the function for the SVM. The comparison reveals that SVM's results are not well-performed, implying that the assumed exponential function in the SVM (eqn (20)) might not well fit the structure-relation of TRSMVs. In contrast, the combinations of 5 linear piecewise functions (ReLU) are employed in the ANN, proving to be a more suitable choice. Second, the number of fitting parameters in the two ML approaches differs. In the SVM, the number of parameters can be calculated as

$$N_\text{p} = N_\text{train}\cdot N_\text{Lag} \tag{24}$$

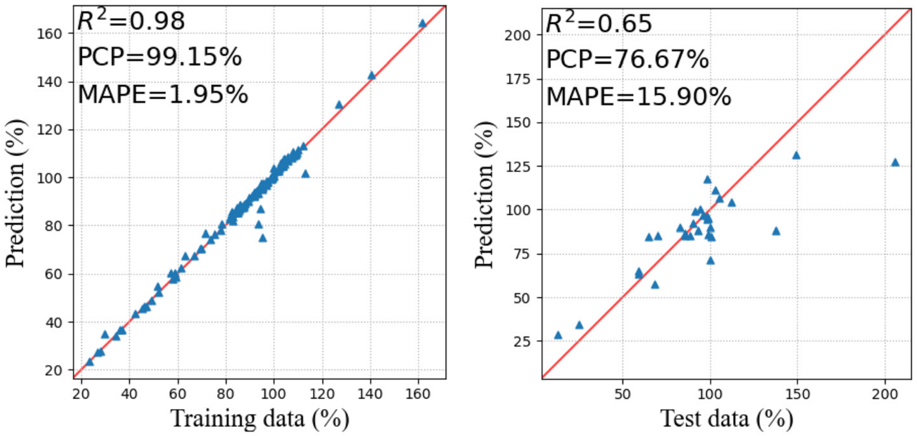
wheren N_train and N_lag are the numbers of training samples for T_g and the numbers of Lagrange multipliers (it can be clearly

Table 2 The comparison of prediction discrepancies among the VAE + ANN model, VAE + SVM, Morgan encoding + SVM, Morgan encoding + ANN

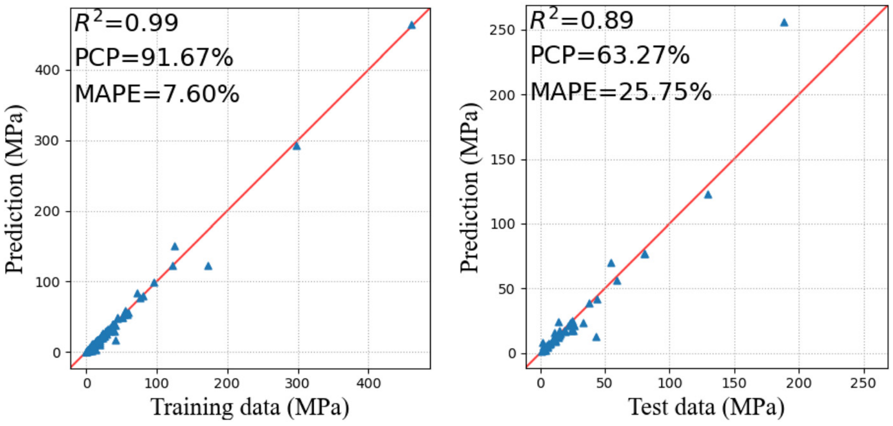
Fingerprinting methods	Supervised learning model	Model output	R^2 in training data	PCP in training (%)	MAPE in training data (%)	R^2 in test data	PCP in test (%)	MAPE in test data (%)
VAE encoding	**VAE + ANN**	$\boldsymbol{M_{T_g}}$	**0.98**	**99.68**	**1.25**	**0.64**	**94.87**	**4.33**
		$\boldsymbol{M_{R_e}}$	**0.98**	**99.15**	**1.95**	**0.65**	**76.67**	**15.90**
		$\boldsymbol{M_{R_m}}$	**0.99**	**91.67**	**7.60**	**0.89**	**63.27**	**25.75**
VAE encoding	VAE + SVR	M_{T_g}	0.76	93.57	5%	0.65	96.15	5
		M_{R_e}	0.56	82.05	16.00	0.19	73.33	38.00
		M_{R_m}	0.33	61.98	35	0.52	42.86	44.00
Morgan encoding	Morgan encoding + SVR	M_{T_g}	0.81	94.53	4	0.43	94.87	5
		M_{R_e}	0.64	83.76	14.00	0.26	70.00	37.00
		M_{R_m}	0.3	65.10	31.00	0.48	36.73	47.00
Morgan encoding	Morgan encoding + ANN	M_{T_g}	0.96	99.68	2.06	0.54	92.31	5.55
		M_{R_e}	0.84	97.44	7.59	0.57	63.33	16.80
		M_{R_m}	0.99	96.35	6.82	0.95	61.22	29.81



(a) T_g training data (b) T_g test data



(c) R_e training data (d) R_e test data



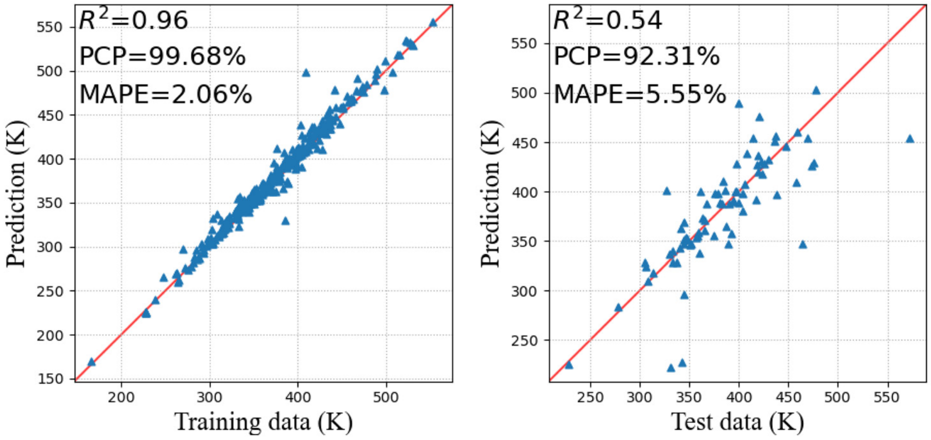
(e) R_m training data (f) R_m test data

Fig. 8 Visualization comparisons for T_g, R_e and R_m based on the VAE + ANN model.

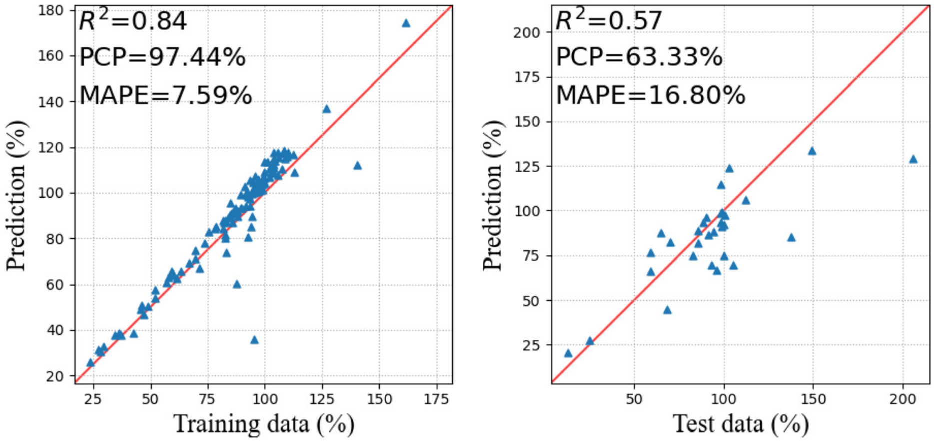
seen that $N_{\text{lag}} = 2$ in eqn (22)). In the ANN, the number of parameters can be calculated using

$$N_p = \sum_{2}^{n} N_{\text{neo}_n} \cdot N_{\text{neo}_{n+1}} + N_{\text{neo}_n} \tag{25}$$

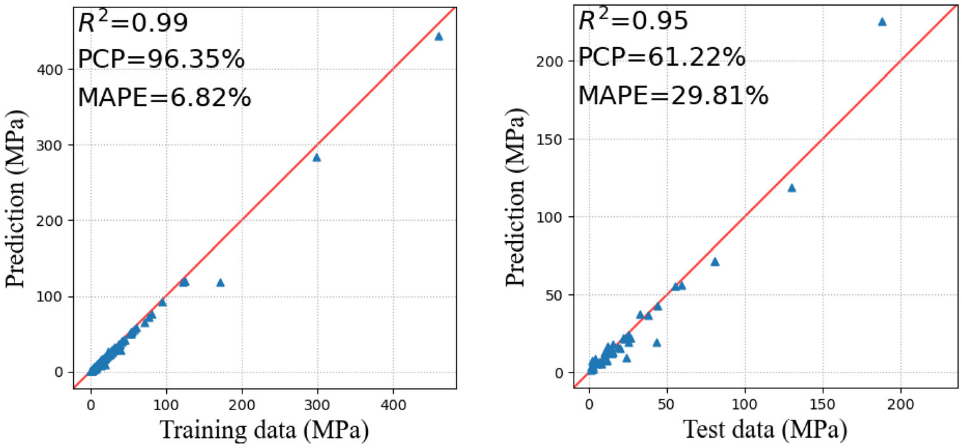
where $N_{\text{neo}_n} \cdot N_{\text{neo}_{n+1}}$ calculates the number of weights from the $(n - 1)$th layer to the nth layer. N_{neo_n} represents the number of biases in the nth layer. Accordingly, the number can only reach up to 389 × 2 = 778 in the SVM, while there are a total of 2 909 633 parameters in the ANN. As polymer features are



(a) T_g training data (b) T_g test data



(c) R_e training data (d) R_e test data



(e) R_m training data (f) R_m test data

Fig. 9 Visualization comparisons for T_g, R_e and R_m based on the Morgan + ANN model.

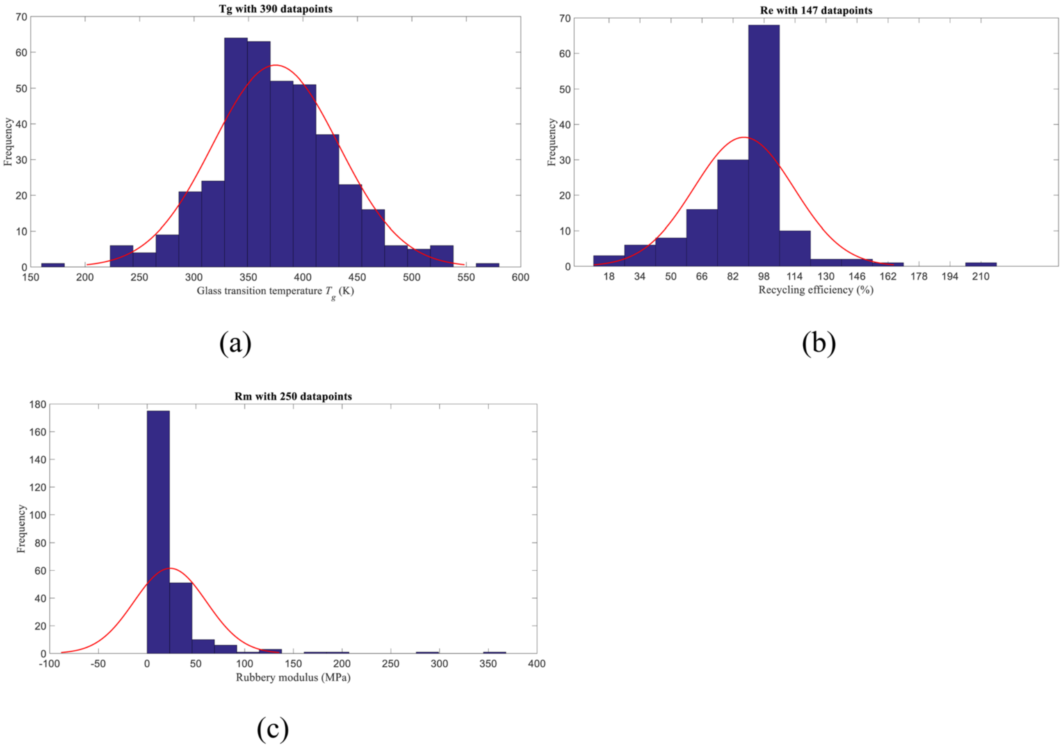
relatively complex, more parameters are essentially needed to fit. This further enhances the results of the ANN model. Among them, VAE + ANN possesses a slight advantage. It means that TL-VAE does extract some hidden structural features from the initial SMILES notation, including the number of rigid chains, number of soft chains and the influences of different fragments, *etc.*

Consequently, our model proves to be a viable and effective option and will work as the desired ML model for subsequent virtual screening. In addition, the other reason why we choose TL-VAE is that this model has the potential to produce new monomers or crosslinkers, which has been validated in our previous research.[35]

Among the models for the three different design targets, the model for T_g exhibits superior performance over the other two models, which can be partially understood by the larger number of data points and close to Gaussian data distributions. Evidently, the dataset of T_g has more data points than the other two datasets, thus allowing the ML model capture more underlying patterns and relationships in the datasets. Furthermore, as illustrated in Fig. 10, the distribution of T_g closely resembles a Gaussian distribution, while the distributions of the other two datasets show neither Gaussian distribution nor uniform distribution. As a result, the data points for T_g may possess better generalization capabilities. These discrepancies could arise from the current research trend, which focuses on high-recycling efficiency vitrimers or extremely softer polymers. Consequently, more extreme values emerge, increasing the likelihood of overfitting and reducing generalization for the models with recycling efficiency and rubbery modulus. It should be noted that we did not implement cross validation in this work. That is, our datasets, comprised of 390, 147, and 250 data points for the three respective datasets, posed unique challenges that rendered traditional methods like cross-validation less suitable. One primary concern was the skewed distribution within our datasets, which could severely impact the predictive performance of our model if cross-validation was employed. For instance, due to the skewed selection of academic journals (only excellent results have the chance to be reported), only a scant number of SMPs can achieve a rubbery modulus exceeding 220 MPa as depicted in Fig. 1 of the manuscript. Should these entries be excluded from the training data, our model might struggle to identify vitrimers with high recovery efficiency owing to the inherent fitting nature of our machine learning models. This scenario could lead to conventional cross-validation to magnify biases, detrimentally affecting predictive performance in the chemical space. To address this, we adopted an alternative approach. We explored 50 distinct random states from 1 to 50, aiming to find an optimal data split that ensures model performance while controlling overfitting. This method, although unconventional, was deemed necessary due to the constraints of our data collection methodology. Our results, as presented in Table 2, support our approach. The variance in MAPE between training and testing sets for all models is consistently below 20%. When this is considered in conjunction with two other performance metrics, we observe only minimal signs of overfitting. Thus, we are confident that our chosen method provides a balanced and pragmatic solution to the challenges inherent in our datasets.

2.5 New TRSMV discovery

2.5.1 Chemical space creation. In order to discover new TRSMVs, we initially created a vast chemical space and then screened for the desired polymers within it. First, we gathered 184 crosslinkers and monomers from our dataset, which



(a) (b)

(c)

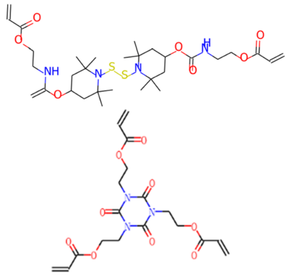
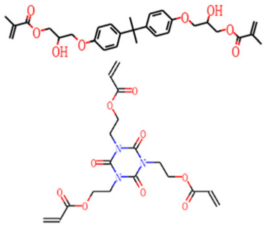
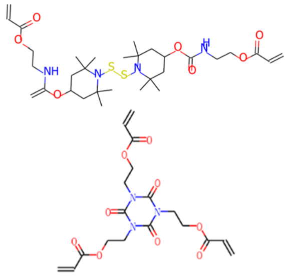
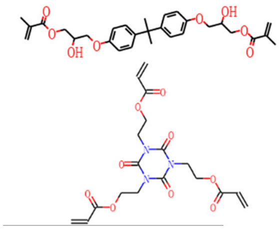
Fig. 10 Histograms for T_g, recycling efficiency and rubbery modulus in the training datasets, which consists of 389, 147 and 250 data points, respectively.

consisted of 9 alcohol molecules (–OH), 30 carboxylic acid molecules (–COOH), 38 epoxy molecules (–OH_3C_2), 55 amine molecules (–NH_2), 11 thiol molecules (–SH), and 41 alkene molecules (–C═C). This chemical space can produce virtual thermosetting polymers synthesized through five classic chemical reactions, including: (1) epoxy-amine reaction, (2) epoxy-hydroxy group reaction, (3) thiol–ene click reaction (C═C reacting with –SH), (4) hydroamination (C═C reacting with –NH_2), and (5) dimerization (C═C reacting with C═C). These five reaction types result in 7959 virtual thermosetting polymers. To decrease the computational workload, we restricted our space to dimers synthesized from only two monomers.

We assume that a total of 1 mol molecules participates in the reaction. By varying the molar ratio of the two monomers between 0.1 : 0.9 and 0.9 : 0.1, we generated a vast chemical space containing 71 631 virtual thermosetting polymers. Since our training dataset only includes 389 polymers, at least 99.45% of the polymers in the chemical space are new polymers. Furthermore, to accurately calculate the molecules participating in the reactions, we introduce an effective stoichiometric ratio to model the varying crosslink densities. For instance, for epoxy (EPON 826) cured by a rigid isophorone diamine (IPD), given the molar ratios 0.1 : 0.9, the effective stoichiometric ratios are 0.1 : 0.05, which is different from the initial molar ratio provided. Apparently, the actual chemical reactions also involve experimental conditions and catalysts; however, we have chosen to discard these factors and assumed that all chemical reactions can fully occur. This comprehensive chemical space can be explored through virtual screening to identify promising polymers with desired properties.

2.5.2 Virtual screening. Unavoidably, our model prediction has some discrepancy. To further improve the screening accuracy, we applied three screening criteria: threshold value screening, dynamic bond screening, and flexible chain screening. First, three threshold values are established for initial screening: $E_r > 250$ MPa, $R_e > 70\%$, and $T_g > 473$K. Given that the highest T_g in our database is approximately 572 K and

Table 3 ML discovered TRSMVs and their predicted target properties, as well as experimental validation for one TRSMV

No.	Combination	Chemical structures of monomers		T_g (K)	E_r (MPa)	R_e (%)
1	HMA : TAI = 0.2 : 0.8		ML	518.43	352.27	75.72
2	BIS-GMA : TAI = 0.2 : 0.8		ML Expt.	520.51 506.5	374.28 332.2	72.83 81.4
3	HMA : TAI = 0.3 : 0.7		ML	503.71	259.39	78.55
4	BIS-GMA : TAI = 0.4 : 0.6		ML	492.82	348.68	73.19

no crosslinker has the functionality exceeding three, we anticipate that the highest T_g of our new TRSMVs will not significantly exceed this value, setting $T_g < 600$ K. Secondly, while all the combinations of monomers and crosslinkers can be synthesized as thermosetting polymers, they may not necessarily have the ability to possess the recyclability. Therefore, we performed further screening based on the presence of dynamic bonds. We counted the number of dynamic bond groups in each thermosetting polymer, selecting only those that contained at least one carboxyl group (–COOH) and one hydroxyl functional group (–OH), or at least one imine functional group (C═N), or at least one disulfide bridge (S–S). This ensures that all selected polymers can form esters, carbonate, urethanes, or imines, or disulfides and related groups. Thirdly, to guarantee better structural rigidity in the screened polymers, we eliminate any polymers containing long flexible chains. Specifically, if a monomer or crosslinker had more than 20 flexible chains, they were eliminated from the target materials. This screening process ensures that the virtual screening results are reliable. Ultimately, we obtained four potential TRSMVs, their corresponding monomers and crosslinkers as well as their three targets, which are listed in Table 3. It can be seen that all the four candidates use the triazine ring as a crosslinker, which is typically a rigid ring with a robust thermostable structure; hence the discovered SMVs should process excellent rubbery modulus and high T_g. Meanwhile, given that candidates 1 and 2 have a higher percentage of triazine rings, they should be more satisfied with requirements of thermally robust shape memory vitrimers (TRSMVs). Furthermore, the monomer BIS-GMA (a type of BPA) in candidate 2 should have better strength and thermostability than the bicyclic compound structure in TAI in candidate 1. However, it is worth noting that the presence of rigid chains (triazine ring and BPA) could potentially impact their recycling efficiency. That is, rigid chains might reduce the mobility of polymer chains, making it harder for the material to flow during remolding or reprocessing steps in the recycling process. This challenge could be mitigated through strategic adjustments in the recycling experiments, such as optimizing processing temperature, employing isothermal curing, reducing the size of waste material, *etc.* In summary, candidate 2 stands out as a particularly promising type of TRSMV, harmonizing robust thermal properties with the shape memory effect and the potential for efficient recycling through methodical experimental design.

2.5.3 **Experimental validation.** In order to further validate the predictability of the ML framework, we synthesized the second new TRSMV listed in Table 3. The polymer can be produced by two pairs of monomers (crosslinkers). One is bisphenol A glycerol dimethacrylate (BPAGMA) cured by a photoinitiater, 2-hydroxy-2-methyl-propiophenone (HMP), purchased from Sigma-Aldrich. The other was synthesized by curing the tris[2-(acryloyloxy)ethyl] isocyanurate monomer (TAI) with a photo-initiator, diphenyl(2,4,6-trimethylbenzoyl)phosphine oxide (TPO), also purchased from Sigma-Aldrich. To prepare the polymer network, we began by mixing 50 g of the BPAGMA monomer with 1.5 g of the photoinitiator HMP in a 250 mL beaker and preheated it at 70 °C for 1 hour to decrease the viscosity and allow efficient mixing. Next, we added 200 g of TAI and 15.05 g of the TPO photo initiator to the BPAGMA/HMP mixture, resulting in a BPAGMA : TAI weight ratio of 20 : 80. The mixture was further stirred using a magnetic stirrer at 90 °C for a duration of 2 hours. It was cured in a UV chamber (IntelliRay 600, Uvitron International, USA) for 40 s under 50% irradiation intensity (232 nm, $\sim$45 mW cm^{-2}) followed by 3 hours of thermal curing at 220 °C.

The thermal behaviors were measured using DMA, where the glass transition could be identified by the peak of the tan delta curve, and the rubbery modulus was determined by examining the plateau region in the respective material's mechanical response. The T_g was first determined using Q800 DMA obtained from TA Instruments (New Castle, DE). The specimens used had a dimension of 4.91 mm wide, 3.16 mm thick, and an effective length of 12.46 mm. The test was conducted under the multi-frequency/strain test mode. The frequency was set to 1 Hz, and the amplitude was selected to be 20 μm. During the temperature scan, we first equilibrated the chamber and specimen to −30 °C using liquid nitrogen and kept at this temperature isothermally for 5 minutes. After this, we started collecting data by ramping at 3 °C min^{-1} to 300 °C. The temperature corresponding to the peak of the tanδ was defined as T_g. From Fig. S2 in the ESI,† the T_g was determined to be 233.5 °C. A DSC 4000 calorimeter made by PerkinElmer (Waltham, MA) was also used to determine the T_g. Samples of approximately 4.30 mg were scanned from −50 °C to 350 °C at 10 °C min^{-1}. The tests were performed in a nitrogen environment with a gas flow rate of 20 mL min^{-1}. The second heating cycle was used to determine the T_g to remove the effect of thermal history. From Fig. S3 in the ESI,† the T_g was determined to be 244.72 °C.

The recycling/healing efficiency of the synthesized polymer was determined *via* the tensile test. We first tested the post-cured specimens with the dimension of 59.98 mm × 5.70 mm × 1.44 mm until fracture. Then the fractured specimens were ball milled using a machine (PQ-N2 Planetary Ball Mill, Across International, New Jersey, USA) for 20 hours. The milled powders were gently introduced into a steel mold and recycled in the eXpert 2610 MTS (ADMET, Norwood, MA, USA) chamber at a pressure of 13 MPa for 2 hours at 220 °C. After demolding, rectangular specimens with dimensions 59.98 mm × 4.95 × 3.52 mm were obtained. The recycling efficiency was calculated using the equation provided below

$$R_e = \frac{\sigma_r}{\sigma_0} \times 100\% \tag{26}$$

where σ_r is the tensile strength of the recycled specimen and σ_0 is the tensile strength of the initial specimen without damage. All the tests were repeated 3 times to obtain the mean value and standard deviation. As given in Table 3, the T_g, rubbery modulus and recycling efficiency from experimental measurements are 506.5 K (or 233.5 °C), 332.2 MPa and 81.4%, respectively. The absolute percentage differences between ML predictions and experimental measurements are 2.77%, 12.67% and 10.52%, respectively.

We also tested the shape memory performance of the new TRSMV. The shape memory effect of the discovered BPAGMA-TAI polymer was determined by first heating the MTS chamber for 1 hour at 220 °C to account for thermal expansion of the metal fixtures and then loading cylindrical samples with dimensions of 5.75 mm diameter and 8.20 mm height at a loading rate of 0.5 mm min^{-1} to achieve a compression programming strain of 24%. The chamber was then cooled down to room temperature to fix the programmed shape. Both stress recovery and free shape recovery tests were conducted. The recovery stress was estimated by preheating the MTS chamber at 220 °C to eliminate the thermal expansion of the fixtures, and the programed cylinder of height 6.2 mm and diameter of 7.2 mm was quickly constrained in the fixtures. The recovery stress as a function of time was recorded as stress *versus* time curve, with recovery stress measured at approximately 33 MPa.

In the free shape recovery test, the compression programmed specimens were placed in the heating chamber at 220 °C for 30 minutes, and the recovered height was measured. The shape fixity ratio F and shape recovery ratio *R* were calculated using the equations provided below:

$$F = \frac{\varepsilon_\text{r}}{\varepsilon_\text{l}} \times 100\%, \quad R = \frac{\varepsilon_\text{f} - \varepsilon_\text{r}}{\varepsilon_\text{f}} \times 100\% \tag{27}$$

where ε_f is the fixed strain after load removal, ε_l is the strain before removing the load, and ε_r is the residual strain after free shape recovery. The tests were repeated more than 3 times to get the mean value and standard deviation.

Finally, we obtain the shape fixity ratio and shape recovery ratio as 99.24% and 98.17%, respectively. In a nutshell, the discovered new TRSMV possesses high shape memory performance and high recycling efficiency, validating the superiority of our ML framework. The experimental measurement shows that the recovery stress is about 33 MPa under a compression programming strain of 24%. This shows that this polymer possesses a recovery stress almost equal to the highest record of 35.3 MPa in the previous study.[10] Overall, this TRSMV has high glass transition temperature, high healing/recycling effciency, and excellent shape memory effect, satisfying the desired goals in this study.

3. Conclusion

In summary, relying on ML techniques, we developed a feasible framework to design TRSMVs for overcoming the conflicting requirements between molecular chain mobility and the formation and reaction of dynamic covalent bonds. The framework allows us to virtually screen the TRSMVs from a vast chemical space, and the designed TRSMVs are able to showcase high recycling efficiency, high T_g, and excellent shape memory effects. The framework exhibits great potential to design new polymers during a short period of time. It is expected that this framework can work as a useful tool for polymer scientists to design new polymers. Simultaneously, we acknowledge that the proposed framework still has certain limitations, such as limited accuracy in rubbery modulus prediction due to the small training datasets and the absence of suitable property descriptors. In the future, we aim to gather more datapoints into our dataset and integrate the property descriptor into our model for further improvement.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work is supported by the Faculty Fellow Program (FFP) + Summer Undergraduate Research Funding (SURF) sponsored by NASA and the Louisiana Board of Regents under contract number NASA(2023)-SURF-01, the US National Science Foundation under grant number OIA-1946231 and the Louisiana Board of Regents for the Louisiana Materials Design Alliance (LAMDA), and the US National Science Foundation under grant number HRD-1736136.

References

1 N. J. Van Zee and R. Nicolaÿ, Vitrimers: Permanently cross-linked polymers with dynamic network topology, *Prog. Polym. Sci.*, 2020, **104**, 101233.

2 D. Montarnal, M. Capelot, F. Tournilhac and L. Leibler, Silica-like malleable materials from permanent organic networks, *Science*, 2011, **334**, 965–968.

3 K. Deng, X. Feng, H. Yang and C. Yan, The Rise of Phosphate Ester Exchange in Developing Dynamic Covalent Networks: Advances and Challenges, *Eur. Polym. J.*, 2023, 112286.

4 X. Feng, X. Lin, K. Deng, H. Yang and C. Yan, Facile Ball Milling Preparation of Flame-Retardant Polymer Materials: An Overview, *Molecules*, 2023, **28**, 5090.

5 M. Guerre, C. Taplan, J. M. Winne and F. E. Du Prez, Vitrimers: directing chemical reactivity to control material properties, *Chem. Sci.*, 2020, **11**, 4855–4870.

6 W. Denissen, J. M. Winne and F. E. Du Prez, Vitrimers: Permanent organic networks with glass-like fluidity, *Chem. Sci.*, 2016, **7**, 30–38.

7 G. Oliveux, L. O. Dandy and G. A. Leeke, Current status of recycling of fibre reinforced polymers: Review of technologies, reuse and resulting properties, *Prog. Mater. Sci.*, 2015, **72**, 61–99.

8 A. R. Rahimi and J. M. Garciá, Chemical recycling of waste plastics for new materials production, *Nat. Rev. Chem.*, 2017, **1**, 1–11.

9 L. Lu, J. Fan and G. Li, Intrinsic healable and recyclable thermoset epoxy based on shape memory effect and transesterification reaction, *Polymer*, 2016, **105**, 10–18.

10 A. Li, J. Fan and G. Li, Recyclable thermoset shape memory polymers with high stress and energy output: Via facile UV-curing, *J. Mater. Chem. A*, 2018, **6**, 11479–11487.

11 Q. Shi, K. Yu, X. Kuang, X. Mu, C. K. Dunn, M. L. Dunn, T. Wang and H. Jerry Qi, Recyclable 3D printing of vitrimer epoxy, *Mater. Horizons*, 2017, **4**, 598–607.

12 J. L. Meyer, M. Bakir, P. Lan, J. Economy, I. Jasiuk, G. Bonhomme and A. A. Polycarpou, Reversible Bonding of Aromatic Thermosetting Copolyesters for In-Space Assembly, *Macromol. Mater. Eng.*, 2019, **304**, 1–9.

13 X. Feng, J. Fan, A. Li and G. Li, Multireusable Thermoset with Anomalous Flame-Triggered Shape Memory Effect, *ACS Appl. Mater. Interfaces*, 2019, **11**, 16075–16086.

14 X. Feng and G. Li, High-temperature shape memory photopolymer with intrinsic flame retardancy and record-high recovery stress, *Appl. Mater. Today*, 2021, **23**, 101056.

15 X. Feng and G. Li, Versatile Phosphate Diester-Based Flame Retardant Vitrimers via Catalyst-Free Mixed Transesterification, *ACS Appl. Mater. Interfaces*, 2020, **12**, 57486–57496.

16 B. Li, G. Zhu, Y. Hao and T. Ren, An investigation on the performance of epoxy vitrimers based on disulfide bond, *J. Appl. Polym. Sci.*, 2022, **139**, 1–11.

17 K. Liang, G. Zhang, J. Zhao, L. Shi, J. Cheng and J. Zhang, Malleable, Recyclable, and Robust Poly(amide-imine) Vitrimers Prepared through a Green Polymerization Process, *ACS Sustainable Chem. Eng.*, 2021, **9**, 5673–5683.

18 Z. Song, Z. Wang and S. Cai, Mechanics of vitrimer with hybrid networks, *Mech. Mater.*, 2021, **153**, 103687.

19 H. Memon, Y. Wei and C. Zhu, Correlating the thermomechanical properties of a novel bio-based epoxy vitrimer with its crosslink density, *Mater. Today Commun.*, 2021, **29**, 102814.

20 L. Ning, L. Yuan, G. Liang and A. Gu, Thermally resistant and strong remoldable triple-shape memory thermosets based on bismaleimide with transesterification, *J. Mater. Sci.*, 2021, **56**, 3623–3637.

21 X. Lan, Y. Liu, H. Lv, X. Wang, J. Leng and S. Du, Fiber reinforced shape-memory polymer composite and its application in a deployable hinge, *Smart Mater. Struct.*, 2009, **18**, 024002, DOI: 10.1088/0964-1726/18/2/024002.

22 G. Li and N. Uppu, Shape memory polymer based self-healing syntactic foam: 3-D confined thermomechanical characterization, *Compos. Sci. Technol.*, 2010, **70**, 1419–1427.

23 L. Santos, A. Dahi Taleghani and G. Li, Expandable proppants to moderate production drop in hydraulically fractured wells, *J. Nat. Gas Sci. Eng.*, 2018, **55**, 182–190.

24 A. Mansour, A. Dahi Taleghani, S. Salehi, G. Li and C. Ezeakacha, Smart lost circulation materials for productive zones, *J. Pet. Explor. Prod. Technol.*, 2019, **9**, 281–296.

25 Vernon, B. Lester and H. M. Vernon, Process of manufacturing articles of thermoplastic synthetic resins. *US Pat.*, 2234993, Mar 18, 1941.

26 J. Curie and P. Curie, Development, via compression, of electric polarization in hemihedral crystals with inclined faces, *Bull. Soc. Minerol. Fr.*, 1880, **3**, 90–93.

27 C. Yan and G. Li, The Rise of Machine Learning in Polymer Discovery, *Adv. Intell. Syst.*, 2022, 2200243.

28 P. Nourian, C. D. Wick, G. Li and A. J. Peters, Correlation between cyclic topology and shape memory properties of an amine-based thermoset shape memory polymer: A coarse-grained molecular dynamics study, *Smart Mater. Struct.*, 2022, **31**(10), 105014, DOI: 10.1088/1361-665X/ac8bb5.

29 C. D. Wick, A. J. Peters and G. Li, Quantifying the contributions of energy storage in a thermoset shape memory polymer with high stress recovery: A molecular dynamics study, *Polymer*, 2021, **213**, 123319.

30 J. Fan and G. Li, High enthalpy storage thermoset network with giant stress and energy output in rubbery state, *Nat. Commun.*, 2018, **9**, 642.

31 Y. Bian and X. Q. Xie, Generative chemistry: drug discovery with deep learning generative models, *J. Mol. Model.*, 2021, **27**, 71, DOI: 10.1007/s00894-021-04674-8.

32 J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer and S. Zhao, Applications of machine learning in drug discovery and development, *Nat. Rev. Drug Discovery*, 2019, **18**, 463–477.

33 C. Yan and G. Li, *Encyclopedia of Materials: Plastics and Polymers*, Elsevier, 2021, pp. 1–24.

34 C. Yan, X. Feng, C. Wick, A. Peters and G. Li, Machine learning assisted discovery of new thermoset shape memory polymers based on a small training dataset, *Polymer*, 2021, **214**, 123351.

35 C. Yan, X. Feng and G. Li, From drug molecules to thermoset shape memory polymers: a machine learning approach, *ACS Appl. Mater. Interfaces*, 2021, **13**, 60508–60521.

36 C. Yan, X. Lin, X. Feng, H. Yang, P. Mensah and G. Li, Advancing flame retardant prediction: A self-enforcing machine learning approach for small datasets, *Appl. Phys. Lett.*, 2023, **122**, 251902.

37 J. W. Lee, W. B. Park, B. Do Lee, S. Kim, N. H. Goo and K. S. Sohn, Dirty engineering data-driven inverse prediction machine learning model, *Sci. Rep.*, 2020, **10**, 1–14.

38 A. Challapalli, J. Konlan, D. Patel and G. Li, Discovery of Cellular Unit Cells With High Natural Frequency and Energy Absorption Capabilities by an Inverse Machine Learning Framework, *Front. Mech. Eng.*, 2021, **7**, 1–12.

39 A. Challapalli, D. Patel and G. Li, Inverse machine learning framework for optimizing lightweight metamaterials, *Mater. Des.*, 2021, **208**, 109937.

40 V. N. Novikov and E. A. Rössler, Correlation between glass transition temperature and molecular mass in non-polymeric and polymer glass formers, *Polymer*, 2013, **54**, 6987–6991.

41 S. E. Wolf, T. Liu, S. Govind, H. Zhao, G. Huang, A. Zhang, Y. Wu, J. Chin, K. Cheng, E. Salami-Ranjbaran, F. Gao, G. Gao, Y. Jin, Y. Pu, T. G. Toledo, K. Ablajan, P. J. Walsh and Z. Fakhraai, Design of a homologous series of molecular glassformers, *J. Chem. Phys.*, 2021, **155**, 224503, DOI: 10.1063/5.0066410.

42 Y. Zhang and X. Xu, Machine learning glass transition temperature of polymers, *Heliyon*, 2020, **6**, e05055.

43 R. W. Hall and P. G. Wolynes, Intermolecular forces and the glass transition, *J. Phys. Chem. B*, 2008, **112**, 301–312.

44 X. Feng and G. Li, Catalyst-free β-hydroxy phosphate ester exchange for robust fire-proof vitrimers, *Chem. Eng. J.*, 2021, **417**, 129132.

45 P. G. Falireas, J. M. Thomassin and A. Debuigne, Imidazolium-catalyzed dynamic ester cross-links towards reprocessable epoxy vitrimers, *Eur. Polym. J.*, 2021, **147**, 110296.

46 Y. Zhu, F. Gao, J. Zhong, L. Shen and Y. Lin, Renewable castor oil and DL-limonene derived fully bio-based vinylogous urethane vitrimers, *Eur. Polym. J.*, 2020, **135**, 109865.

47 Z. Ma, Y. Wang, J. Zhu, J. Yu and Z. Hu, Bio-based epoxy vitrimers: Reprocessibility, controllable shape memory, and degradability, *J. Polym. Sci., Part A: Polym. Chem.*, 2017, **55**, 1790–1799.

48 H. Goss, The Rise of Machine Learning, *Adv. Intell. Syst.*, 2023, **2200243**, 1–29.

49 D. Rogers and M. Hahn, Extended-Connectivity Fingerprints, *J. Mol. Model.*, 2010, **50**(5), 742–754, DOI: 10.1021/ci100050t.

50 D. Weininger, A. Weininger and J. L. Weininger, SMILES. 2. Algorithm for Generation of Unique SMILES Notation, *J. Chem. Inf. Comput. Sci.*, 1989, **29**, 97–101.

51 V. N. Vapnik, *Statistics for Engineering and Information Science Springer Science + Business Media, LLC*, Springer, New York, 2nd edn, 2000.

52 T. Fletcher, *Support Vector Machines Explained*, 2008.

53 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, É. Duchesnay, D. Cournapeau and M. Perrot, Scikit-learn: Machine Learning in Python, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.