ORIGINAL ARTICLE

Transformer-Based Mechanical Property Prediction for Polymer Matrix Composites

Jaewook Lee[1] · Jinkyung Son[1] · Juri Lim[1] · In Kim[2] · Seonwoo Kim[2] · Namjung Cho[2] · Woojin Choi[2] · Dongil Shin[1]

Abstract

Combinatorial nature of polymer matrix composites design requires a robust predictive model to accurately predict the mechanical properties of polymer composites, thereby reducing the need for extensive and costly trial-and-error approaches in their manufacturing. However, traditional prediction models have been either lacking in accuracy or too resource-intensive for practical use. This study proposes an advanced Transformer-based predictive model simultaneously considering various variables that can influence mechanical properties, while utilizing only a minimal amount of training data. In developing this model, we utilize an extensive dataset across 294 types of polymer composites, using a diverse range of polymers and reinforcements, providing a comprehensive basis for the model's predictions. The model employs a Transformer-based transfer learning technique, known for its efficiency with small datasets, to predict essential mechanical properties such as tensile strength, tensile modulus, flexural strength, flexural modulus and density. It shows high predictive accuracy ($R^2 = 92\%$) and makes reliable predictions for combinations of polymer composites that have not been trained on ($R^2 = 82\%$). Additionally, the model's effectiveness and learning process are validated through Explainable Artificial Intelligence analysis and latent space visualization.

Keywords Artificial intelligence · Mechanical property prediction · Polymer matrix composites · Language model

Introduction

Internationally, energy transition policies are being pursued, and social demands for sustainable energy are increasing. Accordingly, demand in new markets such as renewable energy and electric vehicles will increase. In fact, with the continued interest of the international market, the green mobility market has grown. To solve the problem of green mobility energy efficiency, we need to change vehicle materials [1]. So, the stable acquisition of new materials is important for achieving efficient energy transition. Currently, many parts of automobile materials are being replaced by heavy steel materials with light aluminium alloy/plastic/ceramic. As shown in Table 1, replacing each part of an automobile with plastic can achieve a weight reduction ranging from a minimum of 22.2% to a maximum of 66.7% [2]. Plastic, with the advantages of lightweight, moldability, and low production cost, has a distinct drawback of having low mechanical strengths, so the need to develop polymer matrix composites (PMCs) by supplementing the characteristics of plastics with weak strength is required. The material's processability, temperature-dependent strength, and elasticity are major considerations for the practical application of materials. The stretching force of the material, the section of the force returning to the original shape, the braking force, and the temperature at which the deformation occurs can be observed in the mechanical properties test (tensile/flexural test). The working environment (temperature, the velocity of applied force, etc.) to which it is exposed is essential to PMC design in consideration of mechanical properties. Mechanical properties are strongly influenced by the product's intended use and the environment to which it is exposed, so it is closely related to the safety as well as the performance of the product. And the representative mechanical properties are tensile strength, flexural strength, elastic modulus, stress at break, and so on. As mentioned above, polymers

✉ Dongil Shin
dongil@mju.ac.kr

[1] Department of Chemical Engineering, Myongji University, Yongin 17058, Gyeonggi-Do, Korea

[2] Chemical Materials Solutions Center, Korea Research Institute of Chemical Technology (KRICT), Daejeon 34114, Republic of Korea

Table 1 Weight reduction in automobile parts and the rate of weight reduction [2]

	Parts	Existing material	New material	Rate of weight reduction (%)
Engine	Cylinder head cover	Aluminium alloy	Plastic	39.0
	Fuel tank	Steel	Plastic	38.6
	Fuel hose	Steel	Plastic	66.7
Chassis	Brake pipe	Steel	Plastic	22.2
Body	Loop	Steel	Plastic	26.3

have many advantages, but they have the disadvantage of weak strength. Therefore, strengths, including tensile and flexural strength, are the most important mechanical properties in PMC design. However, even with the same PMC, the mechanical properties change according to the filler content, test conditions, and process variables. So, in designing PMC with required mechanical properties, controlling variability in formulation and test that reflects chemical properties is difficult. Consideration of various parameters is required to develop a PMC suitable for the application. But even today, it is difficult to predict the variability of properties affected by various parameters. Many PMC designs are being made based on the intuition of the designer's experience. However, during material development, a lot of trial and error and invasive tests are made in the testing process to measure mechanical properties, which takes a lot of time and physical costs in the process. Consequently, research is being conducted on predictive models for the mechanical properties of PMCs to enable rapid and accurate material design.

The earliest model developed for predicting the properties of PMCs is the rule of mixtures (ROM) model [3–5]. This method estimates the properties of the composite material based on the properties and proportions of its individual components. It is primarily used for predicting the properties of composite materials like polymer composites and fiber-reinforced plastics. However, ROM is based on the overly simplistic assumption that the properties of each component contribute linearly to the overall properties of the composite. This approach fails to account for the effects of interactions between the components, manufacturing process variables, and testing conditions on the material properties. Therefore, the ROM model can only be used for setting a general direction in the initial design phase through rough estimations. Subsequent research is focused on first-principle based models, primarily utilizing Molecular Dynamics (MD) for the prediction of mechanical properties. While the MD model can reflect test temperature, test speed, and structural information of the constituent materials, it incurs a high computational cost and requires extensive data about the structure and properties of these materials. As evidenced by numerous prior studies [6–8], most analyses have been conducted on only one type of material, making high-throughput screening for PMCs under various conditions and combinations

unfeasible. To overcome these limitations of existing models, recent researchers have proposed machine learning-based models that can simultaneously consider a vast array of parameters and accurately reflect the high non-linearity of PMCs. However, acquiring the necessary data for developing predictive models for the mechanical properties of polymer composites requires going through the stages of material design, manufacturing, and mechanical testing. Therefore, most research has relied on limited data obtained by fixing certain variables. The machine learning-based models developed using such limited data have limitations as outlined in Table 2. Additionally, the limitations of recently proposed predictive models become evident through Table 3 [5, 9–14]. These limitations include the neglect of correlations among variables such as the combinations of constituent materials, conditions of the manufacturing process, and the properties to be predicted. As a result, these models only make predictions within a very narrow range and fail to consider a diverse array of condition variables to be viable for industrial use. In response to these challenges, this study proposes a predictive model for the mechanical properties of PMCs that incorporates the maximum possible range of variables from the design process to mechanical testing. To enhance the predictive capability and generalizability of the predictive model, we have employed Transformer-based transfer learning technique. This strategy combine the strengths of Transformer models, which excel at capturing complex patterns and relationships in large datasets, with the benefits of transfer learning, which leverages pre-trained knowledge from extensive datasets. This combination aims to allow our

Table 2 Limitations of existing AI-based predictive models for the mechanical properties of PMCs

Limitations list	
1	Use of hypothetical data
2	Modelling for only one specific type of polymer
3	Modelling for only one specific type of reinforcement
4	Not reflecting a wide range of compositions
5	No consideration of processing conditions
6	No consideration of testing conditions
7	Prediction of only one mechanical property

Table 3 Related study in predicting mechanical properties of PMCs and limitations

Authors	Title	Limitations
Kaweesa et al. [5]	Prediction and validation of composite mechanical properties resulting from voxel-based microstructural design in material jetting	1, 2, 3, 5, 6
Liu et al. [9]	Machine learning assisted prediction of mechanical properties of graphene/aluminium nanocomposite based on molecular dynamics simulation	1, 2, 3, 5
Ho et al. [10]	Development of artificial intelligence-based model for the prediction of Young's modulus of polymer/carbon-nanotubes composites	1, 3, 5, 6, 7
Zhang et al. [11]	Predicting stress–strain curves using transfer learning: knowledge transfer across polymer composites	2, 3, 4, 6
Gulihonenahali et al. [12]	An artificial neural network prediction on physical, mechanical, and thermal characteristics of giant reed fiber reinforced polyethylene terephthalate composite	3, 5, 6
Al-Jarrah et al. [13]	A novel integrated BPNN/SNN artificial neural network for predicting the mechanical performance of green fibers for better composite manufacturing	4, 5, 6
Iqbal et al. [14]	Prediction of residual tensile strength of glass fiber reinforced polymer bars in harsh alkaline concrete environment using fuzzy metaheuristic models	1, 2, 3, 5

model to consider intricate dependencies and interactions between variables, reduce the need for extensive labelled data, and improve learning efficiency. By leveraging these capabilities, we propose a predictive model that could deliver more accurate predictions and effectively manage the high non-linearity inherent in PMCs.

Data Collection, Analysis and Split for Model Training and Validation

Mechanical Property Data of PMCs

In this study, mechanical property data of PMCs, which are based on polymers and reinforcements most commonly used in lightweight materials, were utilized [2, 15]. The total number and types of data are as shown in Fig. 1. The polymers used for the specimens manufactured for mechanical testing are a total of 8 types: Polypropylene (PP), Polycarbonate (PC), Polyamide6 (PA6), Polyamide46 (PA46), Polyamide66 (PA66), Acrylonitrile butadiene styrene (ABS), Styrene-acrylonitrile (SAN), and Polybutylene terephthalate (PBT). The reinforcements used are a total of 7 types:

Glass fiber (GF), Carbon fiber (CF), Amide fiber (AF), Talc, $CaCO_3$, Ethylene Octene Rubber (EOR), and Boron nitride (BN). The total number of combinations and compositions of the manufactured PMCs is 294. Additionally, 31 process-related variables were identified for each material, and the compositions' distribution is based on mass fractions of 0%, 5%, 10%, 15%, 20%, 30%, 40%.

Tensile tests were conducted using a Universal Testing Machine (UTM), and the tests were conducted in adherence to the globally recognized standard for measuring mechanical properties of plastic materials, ISO 527. The gauge length (L0) was set at 50 mm, the test speed for calculating the modulus of elasticity was 1 mm/min, and the test speed after the modulus of elasticity phase was conducted at 5 and 50 mm/min. During these tests, parameters such as the combination, composition, manufacturing process conditions, and testing conditions of the manufactured PMCs were recorded. The thickness and length of the specimens were uniformly manufactured to be 4 cm and 10.13 cm, respectively, in accordance with standards. To account for external temperature exposure and force speed, the test temperatures were set at − 30 °C, 23 °C, 80 °C, and 120 °C, while the tensile speeds were set at 5, 50, and 500 mm/min. Additionally,
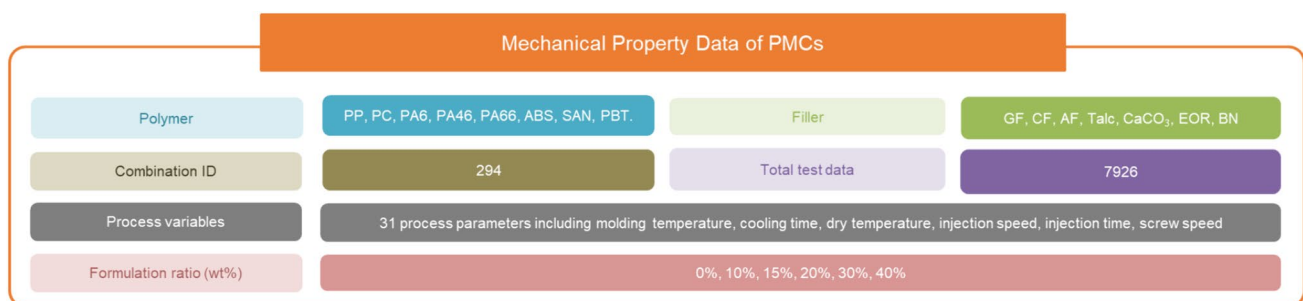
Fig. 1 Acquired mechanical property data of polymer composite materials

the specimens were stored in a constant temperature and humidity room maintained at 23 °C and 50% humidity for over 72 h before testing. The properties obtained included tensile modulus and tensile strength.

The flexural tests were also conducted using the same UTM equipment in a constant temperature and humidity room maintained at 23 °C and 50% humidity. Following the international testing standard ISO 178, these tests were performed with a span distance of 64 mm and a testing speed of 2 mm/min. The properties obtained from these tests were flexural modulus and flexural strength. Additionally, to accurately measure the density of the synthesized materials, a densimeter (EW-300SG, Alfa Mirage) was used in the same constant temperature and humidity room. The specific gravity was measured and then multiplied by the density of water at 23 °C to calculate the density of the specimens. Measurements were taken using flexural specimens, and each composition involved five specimens that were stored

in the constant temperature and humidity room for over 72 h before the data was recorded.

Examination of Correlations Between Mechanical Properties and Process Variables

To begin with, Fig. 2 allows us to conduct a basic analysis of five mechanical properties. The tensile and flexural strengths can be observed in Fig. 2a, c, respectively, with most of the values concentrated below 150 MPa. This indicates that the mechanical properties of the polymer, which is the matrix of the PMC, are significantly reflected. As shown in Fig. 2b, d, the modulus values are primarily distributed between 1000 and 3000 MPa, and Fig. 2f reveals that both the tensile modulus and flexural modulus have considerable dispersion. Lastly, through Fig. 2e, f, the density of PMCs based on 15 different constituent materials is found to vary significantly, ranging from a minimum of 0.888 g/cm^3 to a
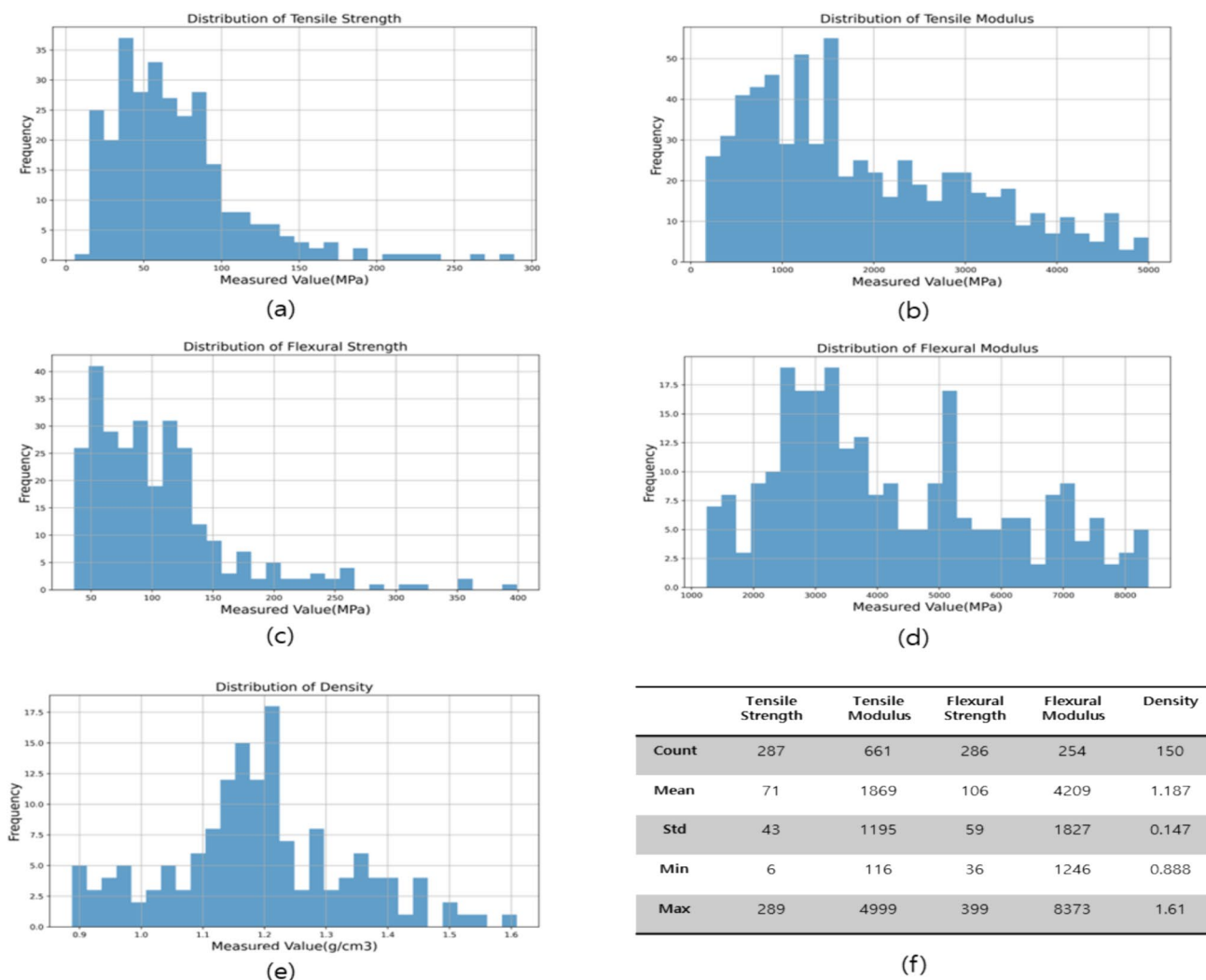

	Tensile Strength	Tensile Modulus	Flexural Strength	Flexural Modulus	Density
Count	287	661	286	254	150
Mean	71	1869	106	4209	1.187
Std	43	1195	59	1827	0.147
Min	6	116	36	1246	0.888
Max	289	4999	399	8373	1.61

(f)

Fig. 2 Analysis results of PMC mechanical properties raw data

maximum of 1.61 g/cm^3, approximately a twofold difference. The material with the highest mechanical strength in the collected data is a PMC manufactured with PA6 and CF as filler at 40 wt%, which exhibits a tensile strength of 344 MPa and a flexural strength of 353 MPa. For lightness of materials, PMCs reinforced with CF and using PA6, PA66, PA46 as the polymer matrix generally showed high mechanical strength. However, without simultaneous analysis of the other 22 variables, including process and testing conditions, these trends cannot be simply accepted. In other words, it is not clear from this simple analysis whether the superior mechanical properties of certain combinations are due to random arrangements and bonds formed during the polymer and material mixing process or improved due to optimal manufacturing conditions. To clearly analyze the correlation between processing conditions and the targeted mechanical properties, a Pearson Correlation Coefficient (PCC) analysis was conducted [16]. The PCC is a statistical technique measuring the strength and direction of a linear relationship between two variables, ranging from − 1 to 1. A PCC of 1 indicates a perfect positive linear relationship,

− 1 indicates a perfect negative linear relationship, and 0 indicates no linear relationship. Generally, a value of ± 0.2 or higher is considered indicative of a correlation between variables [17].

In the PCC analysis for tensile strength, shown in Fig. 3a, 14 out of the 19 process parameters exhibited a PCC value of ± 0.2 or higher. Similarly, for flexural strength, Fig. 3c reveals that 15 out of 19 process parameters have a PCC value of ± 0.2 or higher. These results clearly demonstrate the significant relation and collinearity between the mechanical properties of PMCs and the manufacturing process variables obtained. This indicates that manufacturing process variables can be effectively utilized as input features in predicting the mechanical properties of PMCs. However, as shown in Fig. 3b, d, there is no clear correlation between the tensile modulus, flexural modulus, and the process parameters. This suggests that the modulus of elasticity, which is primarily related to the basic bonding characteristics of the material, might be less influenced by the stress distribution within the material during the extrusion process compared to its impact on strength. It is important to note that Pearson's
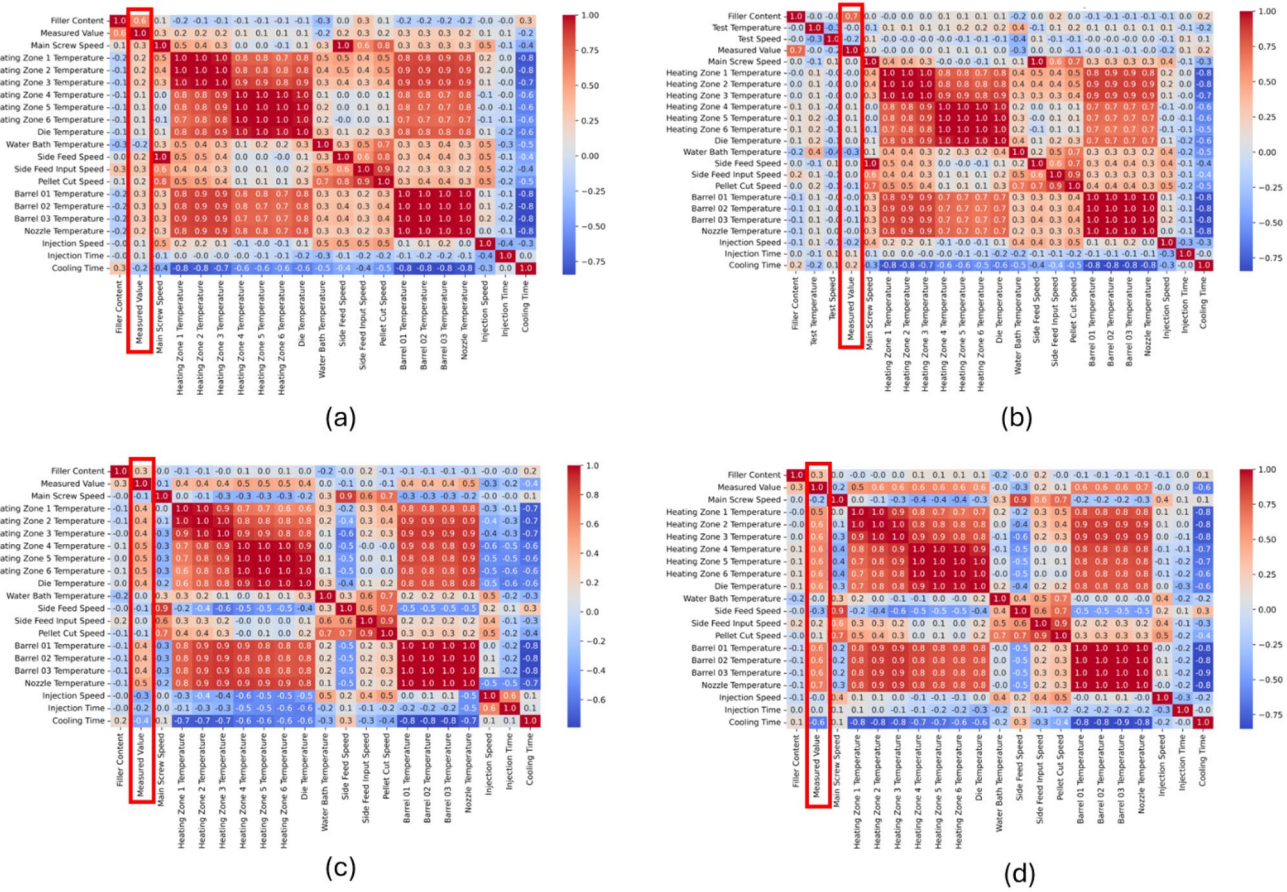

Fig. 3 Analysis results of correlation between mechanical properties and processing conditions. **a** PCC analysis between tensile strength and variables. **b** PCC analysis between tensile modulus and variables. **c** PCC analysis between flexural strength and variables. **d** PCC analysis between flexural modulus and variables

Correlation Coefficient (PCC) analysis only identifies linear relationships, and thus cannot detect higher order nonlinear correlations. Further verification of this aspect was conducted in Sect. 4 through feature engineering, as part of the predictive result analysis process.

Data Preparation for Model Performance Validation

The original dataset contained 31 process-related variables. However, as indicated in Fig. 3, some of these variables exhibited very high correlations with each other. To improve computational efficiency and reduce dimensionality, 9 of these highly correlated variables were replaced with their average values, specifically averaging the temperatures of Heating Zones 1 to 6 and the temperatures of Barrel 1 to Barrel 3. This resulted in a total of 24 input features. The 24 input features and 5 outputs in the collected data exhibit vastly different scales. The data with the largest scale is the flexural modulus, with a scale around 10^4, while the data with the smallest scale is the composition data, with a scale of 10^{-1}, resulting in a maximum scale difference of 10^5. Large differences in scale can cause several issues. Firstly, the disparity in scales can significantly decrease the learning efficiency of machine learning-based models that utilize optimization techniques like gradient descent, potentially slowing down or rendering the learning process inefficient. Moreover, features with larger sizes might have a disproportionately large impact on the results, hindering the proper learning of genuinely important features. Therefore, to maintain the original distribution of the data while scaling all features to a range between 0 and 1, the MinMax scaling technique was applied.

The MinMax scaling technique can be defined by the following Eq. (1), where X_{scaled} represents the adjusted data, and X_{min}, X_{max}, and X are the minimum, maximum, and original values in the dataset, respectively. Utilizing MinMax scaling for range adjustment not only aids in model training but also reduces the impact of outliers and maintains the distribution of data that does not follow a normal distribution.

$$X_{\text{scaled}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}. \tag{1}$$

In this study, two data partitioning methods were adopted to thoroughly verify the reliability and accuracy of the model. First, the entire dataset was randomly divided into training and validation sets in an 8:2 ratio for model training. This approach was taken to assess the predictive accuracy of the model when the composition, process variables, and test variables change within the combinations of trained PMCs.

Second, a method of selectively removing specific datasets from the training data to measure predictive accuracy was used. By excluding the entire data of certain PMC

combinations from the training data and verifying predictive accuracy, the model's performance in predicting untrained PMC combinations was evaluated. The three combinations with the most data were selected from the entire dataset as shown in Table 4 to assess the model's performance.

Methods

Representations and Tokenization of PMCs

In this study, since a natural language processing model is utilized, it is necessary to convert PMCs into SMILES, a natural language form. The Simplified Molecular Input Line Entry System (SMILES) notation, which represents molecular structures as natural language, is the most widely used molecular structure representation and provides richer structural information compared to chemical fingerprints based on general functional groups. To use SMILES representation as input and output for the Transformer model, SMILES strings need to be tokenized into individual characters and then encoded into one-hot vector representations (a vector with all values set to 0 except for the position of the current token). This process is similar to tokenization of sentences in natural language processing, and the key lies in building a vocabulary that accurately represents the structures of polymers and reinforcements. For this purpose, a vocabulary was constructed using the collected polymer and reinforcement data, especially for pre-training in the transfer learning process. By utilizing polymer data, efficient extraction of representations within polymers was possible. The constructed vocabulary consists of tokens for 372 different SMILES structures. Figure 4 illustrates the process of tokenizing polymer SMILES using the constructed vocabulary. When building a vocabulary with general chemical SMILES, tokenization based on elements like carbon (C), hydrogen (H), oxygen (O) leads to a smaller vocabulary size, but the tokenized SMILES become very long, increasing the dimensionality of the data to be interpreted. This can result in increased memory usage for training and longer training times and can also cause sparsity issues for rarely used elements, potentially hindering proper learning. Additionally, including too many words can make the model overly specialized in certain specific words, complicating generalization. However, by constructing a new vocabulary through polymer data, it is possible to effectively reflect structural information of polymers by reducing the size of the vocabulary while converting frequently used SMILES expressions, i.e., functional

Table 4 Polymer composite material combinations for verifying model extrapolation capability

Polymer matrix	Reinforcement
PA46	CF
PP	GF
ABS	GF

Fig. 4 SMILES tokenization result

Symbol	Token
346	[
327	CH
348	]
76	CCCCC
14	N
322	C
1	(
334	=O
296	)
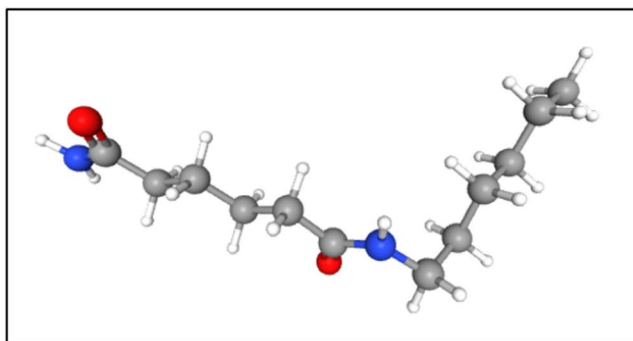
SMILES: [CH]CCCCCNC(=O)CCCCC(=O)[N]
Tokenized SMILES: [346, 327, 348, 76, 14, 322, 1, 334, 296, 76, 1, 334, 296, 346, 14, 348]

groups, into single words. Figure 4 shows the tokenization process for Polyamide 66 (nylon 66), one of the PMCs data used in modelling, demonstrating how CH functional groups or chains like CCCCC are efficiently converted into single words within the tokenization process.

The size of the constructed vocabulary is 372, and it includes specific tokens and indices added for the structure of polymers and reinforcements, as well as other vocabulary, such as: ('<unk>' 368), ('<pad>' 369), ('<bos>' 370), ('<eos>' 371). Among these, <bos> and <eos> are used to denote the beginning and end of SMILES strings, respectively. The <unk> token was added to substitute for expressions not present in the dictionary. Lastly, <pad> is used to fill SMILES strings that are shorter than the model's input size to make their length equal to the input size. Additionally, considering that most of the SMILES representations of polymers and reinforcements in the entire dataset are within 40 characters, the size of the input SMILES was set to 40 characters each.

Design of Transformer-Based Predictive Model Through Transfer Learning

The Transformer model used in this study is based on an encoder-decoder architecture and has been widely used in cheminformatics for predicting molecular properties and reactions [18–21]. Particularly, it has been observed in previous studies that utilizing transfer learning of SMILES shows superior performance compared to other modelling techniques, especially when modelling with a small amount of data. Previous research has demonstrated the success of Transformer-based transfer learning techniques for various molecular properties, such as physical properties (e.g., melting point, solubility), molecular orbital properties (e.g., HOMO, LUMO), and environmental, health, and safety (EHS) characteristics (e.g., mutagenicity, toxicity) [22–24]. Therefore, it was anticipated that the prediction of mechanical properties of PMCs using a pre-trained Transformer model, based on the limited PMC database in this study, could be achieved in a similar manner.

The basic design outline of the model is as shown in Fig. 5, and the model development process consists of pre-training and fine-tuning. The key to this Transformer-based transfer learning framework lies in the pre-training phase. By pre-training on SMILES data, the model can learn the representation of SMILES and efficiently extract complex functional group information, effectively distributing the information represented by SMILES in a latent space. This process allows for the compression of high-dimensional SMILES into a smaller dimension, focusing on learning only the most crucial information, thus increasing the efficiency of the predictive model. Additionally, through pre-training, the Transformer model can learn various patterns and language structures from a larger dataset, enabling it to perform effective learning with less data compared to other machine learning-based models and achieve high performance and reliability.

Since the mechanical property prediction model for PMCs takes SMILES as input, the ability of the Transformer model, a natural language processing model, to accurately interpret the SMILES of the constituent materials is critical to the model's performance and reliability. Therefore, a self-supervised learning approach was used to enable the predictive model to learn SMILES syntax. While it is typical for pre-training databases in chemical SMILES to encompass a wide variety of molecules, in this study, the target materials are polymers and reinforcements, not typical chemical molecules. Therefore, feature extractors obtained from training on general chemical molecules may not guarantee effective extraction of key information from polymers and reinforcements. Consequently, we constructed a pre-training dataset that could effectively learn and extract the structural characteristics of PMCs.

Pre-training with SMILES can be conducted in various ways, with the most commonly used methods being masking-based learning and standardization-based learning.

The first method involves masking specific tokens within the entire SMILES string and then predicting parts of these masked SMILES as a pre-train task, as shown in Fig. 6. The second method involves converting non-canonical SMILES into canonical SMILES, as illustrated in Table 5. Based on the analysis that pre-training by converting non-canonical SMILES to canonical SMILES yields higher performance, this study also selected this approach as the pre-training task [21].

For pre-training, the constructed SMILES database contains information on 8512 polymers and reinforcements. The detailed database construction process began

with collecting the types of polymers and reinforcements from PolyInfo [25], followed by converting their chemical names to SMILES in PubChem to build the database [26]. Subsequently, the pre-training used non-canonical SMILES as input, and the model was trained to predict canonical SMILES as output. Examples of input–output data for this pre-training process are shown in Table 5. The performance improvement of the predictive model through this pre-training process was analyzed in detail in Sect. 4.2. To develop predictive models for tensile strength, tensile modulus, flexural strength, flexural modulus, and density, the pre-training process in Fig. 5a was

Fig. 5 Development flowchart of transfer learning-based mechanical properties prediction model. **a** Unsupervised pre-train. **b** Supervised learning to predict mechanical properties and density of PMC
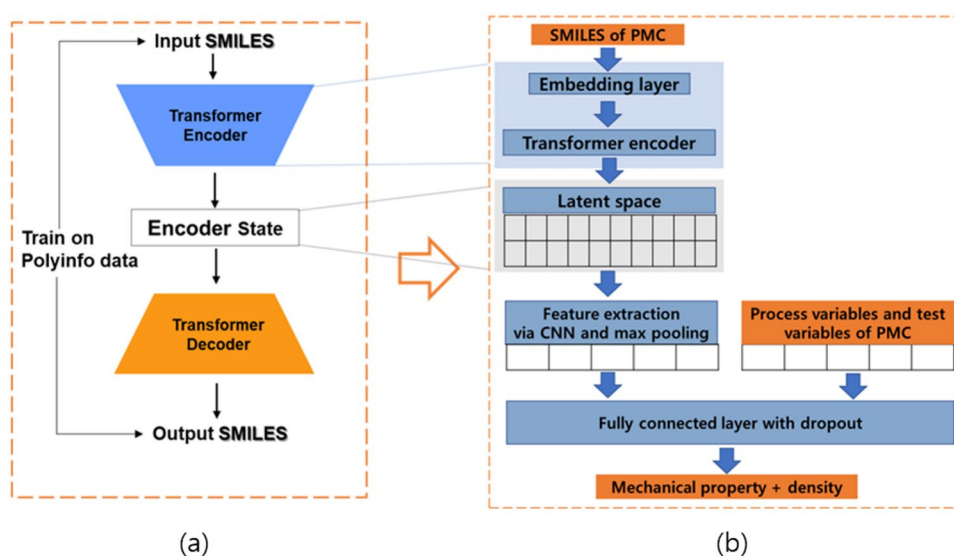

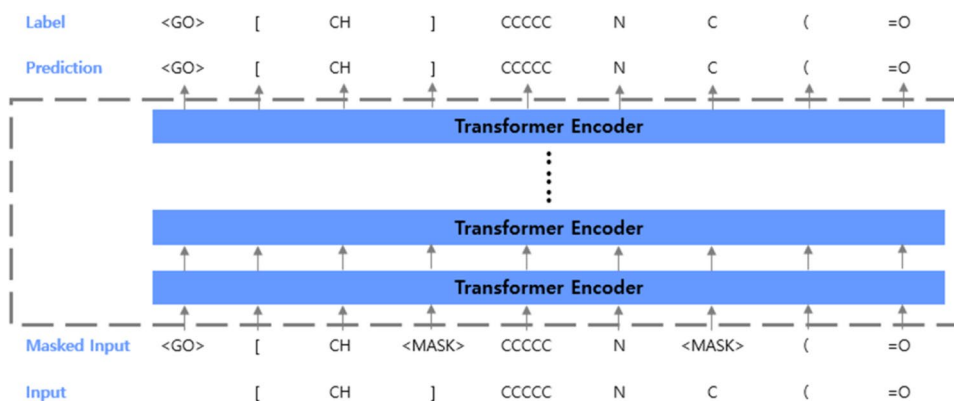**Fig. 6** Masked SMILES pre-training with transformer

Table 5 Input and output examples of pretraining process

Pre-training input (non-canonical SMILES)	Pre-training output (canonical SMILES)
C(CCC[C@@H](C(=O)*)N*)SC©=O	*N[C@@H](CCCCSC©=O)C(*)=O
N[C@H](C(=O))CCCCSC(==O)C	*N[C@@H](CCCCSC©=O)C(*)=O
C(CCC[C@H](N*)C(=O)*)SC(C)=O	*N[C@@H](CCCCSC©=O)C(*)=O
O=C©SCCCC[C@H](N*)C(=O)*	*N[C@@H](CCCCSC©=O)C(*)=O
C(CC[C@@H](C(=O)*)N*)CSC©=O	*N[C@@H](CCCCSC©=O)C(*)=O

used to adjust the parameters of the Transformer encoder, which was then further fine-tuned.

Specifically, the encoder component of the pre-trained Transformer model is used to extract the core features of the input polymers and reinforcements, acting as a molecular fingerprint. For example, for a molecule with a SMILES string of N characters, the encoder generates an N-dimensional latent representation vector. Since different molecules have SMILES of varying lengths, the input size for the downstream model can vary for each polymer and reinforcement SMILES. Therefore, a convolutional neural network (CNN) with Max pooling structure is applied to the downstream predictive model to accommodate the original sentence classification. This design allows the model to accommodate various input lengths and operate regardless of the length of the polymer and reinforcement SMILES representations. However, since the encoder of the pre-trained model only takes polymers and reinforcements SMILES as input, it cannot reflect various input variables such as process parameters and mechanical testing variables. To incorporate these, as shown in Fig. 5b, an additional input layer with a recurrent neural network was attached to develop a multi-input model that reflects the process and testing parameters of PMCs. The overall structure of the model is divided into two parts: one that receives the SMILES of PMCs as input and another that receives additional process and mechanical testing conditions as input. The model structure for the part that receives the PMCs SMILES as input was designed based on the results optimized during the pre-training process, and subsequently, the structure of the feedforward neural network attached as a mechanical property prediction regressor was further optimized through Bayesian optimization.

Results and Discussion

Impact Analysis of Input Features

To validate the performance of the model developed in this study, the first step undertaken was an analysis of the impact of process conditions and mechanical testing conditions on the predictive model. Due to the significant time cost of validating all data, verification was conducted focusing on tensile strength, which had the most data available. For this verification, two models were developed and tested: one that only received PMCs SMILES and composition as input, excluding process and mechanical testing conditions, and another that included all inputs, incorporating process conditions and mechanical test conditions.

The results of the verification can be seen in Fig. 7. Figure 7a shows the tensile strength prediction results of the Transformer model using only PMC SMILES, while Fig. 7b presents the results of the Transformer model that also includes process and testing variables. When only SMILES were used as input, the R^2 value based on the validation data was 63%, and when process and testing variables were also included, the R^2 value increased to 81%, an increase of approximately 18%. This significant improvement demonstrates the importance of process conditions and mechanical testing conditions as vital input features in predicting the mechanical properties of PMCs.

Analysis of Model Performance Changes Through Pre-training

In this part, we analyzed the impact of pre-training on the performance of the model. This analysis was conducted in two ways. First, a predictive model without pre-training was
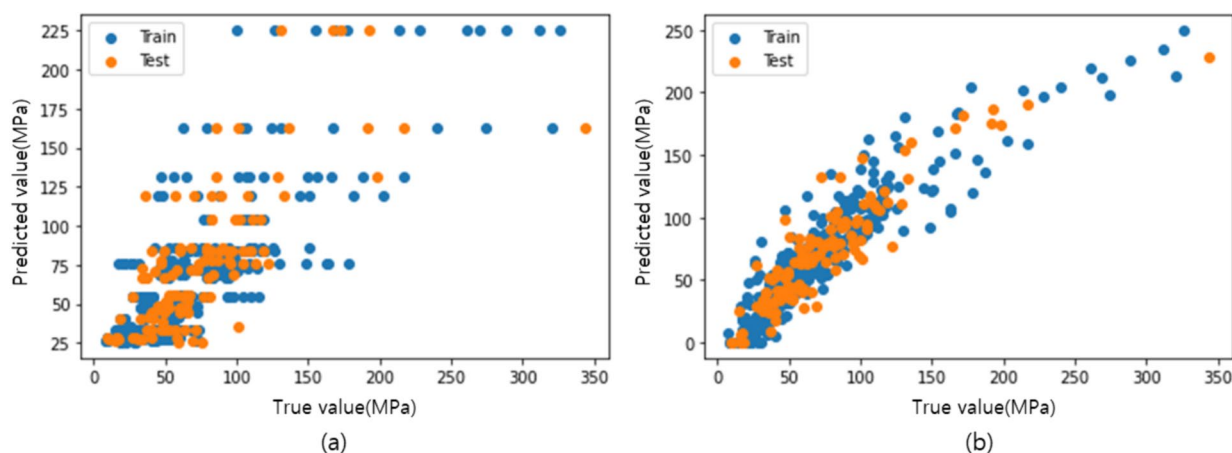

Fig. 7 Tensile strength prediction result comparison with different input features. **a** Using SMILES as input feature (R^2: 63%). **b** Using SMILES, process conditions, test conditions as input feature (R^2: 81%)

developed through supervised learning with mechanical property data, and its predictive performance was compared to that of a model developed with transfer learning. The comparison of prediction results can be seen in Fig. 8. The model trained solely with supervised learning achieved an R^2 value of 81% on the validation data. In contrast, the transfer learning-based model resulted in a significantly higher R^2 value of 95%. This demonstrates a clear performance improvement of approximately 14% with the incorporation of pre-training. This demonstrates the effectiveness of the Transformer predictive model development framework with pre-training in predicting mechanical properties of PMCs.

Second, to prove the performance of the Transformer predictive model developed through pre-training, the latent space of the Transformer model was examined. Both models

with and without pre-training were used, and the latent representation vectors distributed in the latent space when the SMILES of PMCs were input were observed. This involved compressing the high-dimensional 512-dimensional latent representation vectors into a two-dimensional space for visualization, using the uniform manifold approximation and projection (UMAP), which is most effective for embedding SMILES [27]. Figure 9a shows the latent representation vectors of the tensile strength prediction model developed through supervised learning, while Fig. 9b shows those of the model developed with transfer learning. The X and Y axes represent the dimensions of the reduced latent space, and the colors of the data points represent the magnitude of tensile strength. While Fig. 9a shows that mapping to the latent space according to tensile strength is indistinct,
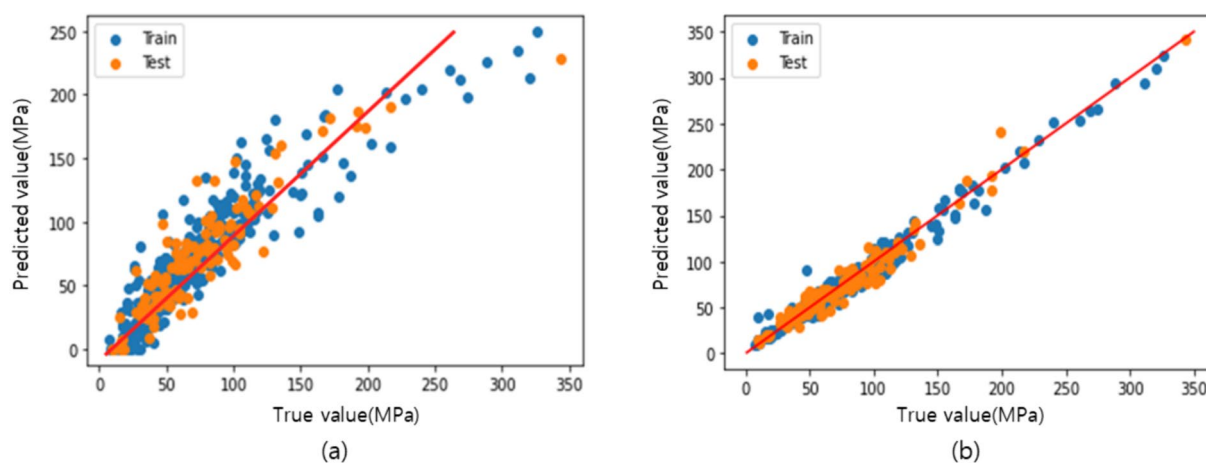

Fig. 8 Tensile strength prediction results. **a** With supervised learning only (R^2: 81%). **b** With transfer learning (R^2: 95%)
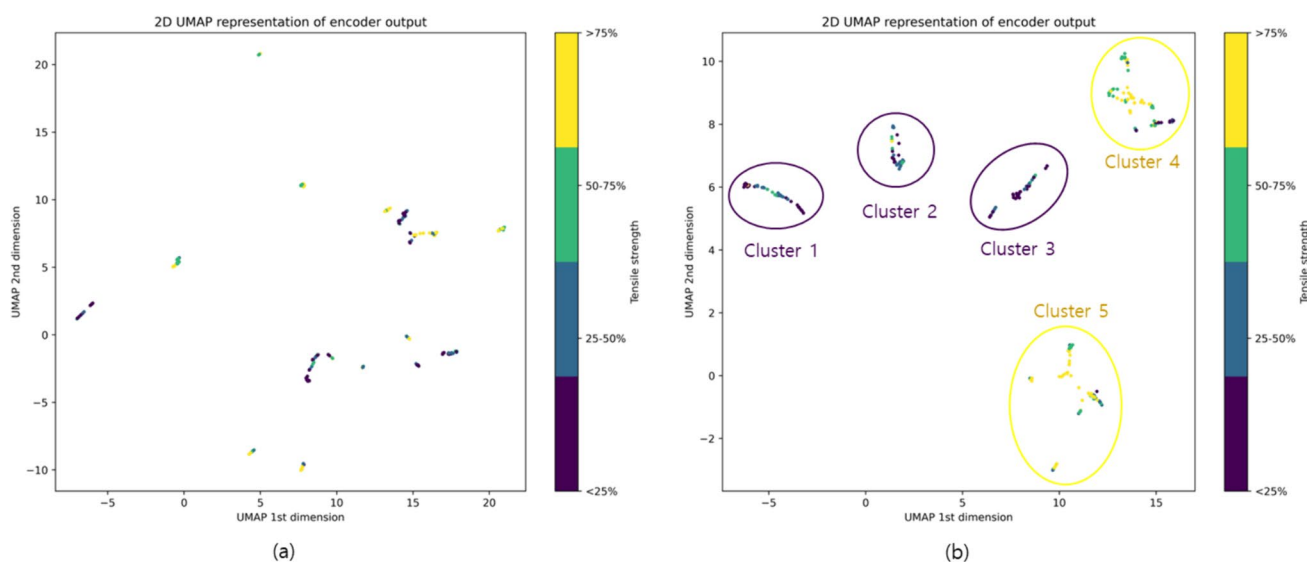

Fig. 9 Latent space of tensile strength prediction model. **a** With supervised learning only. **b** With transfer learning

Fig. 9b clearly shows distinct mappings according to tensile strength. Remarkably, without employing any clustering or classification techniques, five distinct clusters can be observed in accordance with tensile strength, as if clustering based on tensile strength had been performed. This indicates that the encoder of the Transformer, through transfer learning, is more effective in extracting features critical to the mechanical properties of PMCs. Based on the analysis of these two cases and the various transfer learning-based studies discussed in Sect. 3.2, we have determined that transfer learning would have a significant positive impact on the prediction of other mechanical properties as well.

Analysis of Predictive Model Performance by Mechanical Properties

Through the various analyses conducted, the applicability of the model development framework using transfer learning has been validated. Consequently, the results of predicting tensile strength, tensile modulus, flexural strength, flexural modulus, and density using transfer learning were examined. The data were split using the method described earlier, randomly dividing the entire dataset into training and validation sets at a ratio of 8:2 for learning and prediction. The results of the model training can be seen in Fig. 10. The performance of the model was evaluated based on the R^2 value for the validation data. The performances of the predictive models for tensile strength, tensile modulus, flexural strength, flexural modulus, and density were found to be high, with respective R^2 values of 95%, 93%, 90%, 91%, and 92%.

In the problem of Quantitative Structure–Property Relationship (QSPR) for predicting mechanical properties from the molecular structure of chemical substances, the most crucial aspect is the extrapolation ability of the model. To test this, three types of PMCs with the most data among the entire mechanical property dataset were used as validation data to assess the model's extrapolation capability. These types are composites of PA46 and CF, PP and GF, and ABS and GF. All data for these three combinations were excluded from the training process, and their prediction performance was evaluated. The results are shown in Fig. 11, and the prediction performance for these three types of PMCs is 87%, 81%, and 79% based on R^2, respectively. These validation results confirm that the model can robustly predict not only interpolations but also extrapolations for previously unseen combinations of PMCs.

Interpretation of the Predictive Model Through XAI Analysis

The developed mechanical property prediction model is a Transformer-based black-box model, which does not provide an understanding of the criteria for its predictions. To assess the reliability of this model, Explainable AI (XAI) techniques were introduced, and the Local Interpretable Model-agnostic Explanation (LIME) algorithm was employed for model analysis.

LIME is an XAI algorithm that approximates the model near a data point of interest into an interpretable white-box model, providing explanations for the predictions of the black-box model. In this study, the data point with the highest predicted tensile strength was inputted for model interpretation. The analysis results can be seen in Fig. 12. The XAI results indicate that mechanical strength increases with lower test temperatures, slower injection speeds, higher filler composition ratios, and lower side feed speeds for the reinforcements. These insights are consistent with expert knowledge in the field of polymer composite design, thereby validating the reliability of the model. Specifically:

- *Lower Test Temperatures* At lower temperatures, polymer materials exhibit higher stiffness and strength due to reduced mobility of polymer chains. This decreases the thermal motion of molecules, inducing stronger intermolecular interactions.
- *Slower Injection Speeds* A slower injection speed allows the polymer melt to flow more uniformly within the mold, leading to more uniform cooling and crystallization. This minimizes internal stress and forms a more uniform microstructure, improving the material's mechanical properties.
- *Higher Filler Composition Ratios* Reinforcements dispersed within the polymer matrix increase the interaction between polymer chains and help distribute external stress throughout the material, increasing its resistance to load.
- *Lower Side Feed Speeds* A lower feed speed for reinforcements reduces internal stresses during the injection process, leading to more uniform flow of the melt and decreased stress concentration. This reduces the occurrence of microdefects or voids within the material, contributing to the improvement of mechanical properties, especially tensile strength.

By verifying that the model's interpretation aligns accurately with expert knowledge, the reliability of the model is validated. Additionally, the LIME analysis provides clear insights into how specific input variables (e.g., process conditions) affect mechanical properties such as tensile strength. This understanding helps to ensure that the model's predictions are based on plausible and scientifically sound relationships, further enhancing confidence in its predictive capabilities.
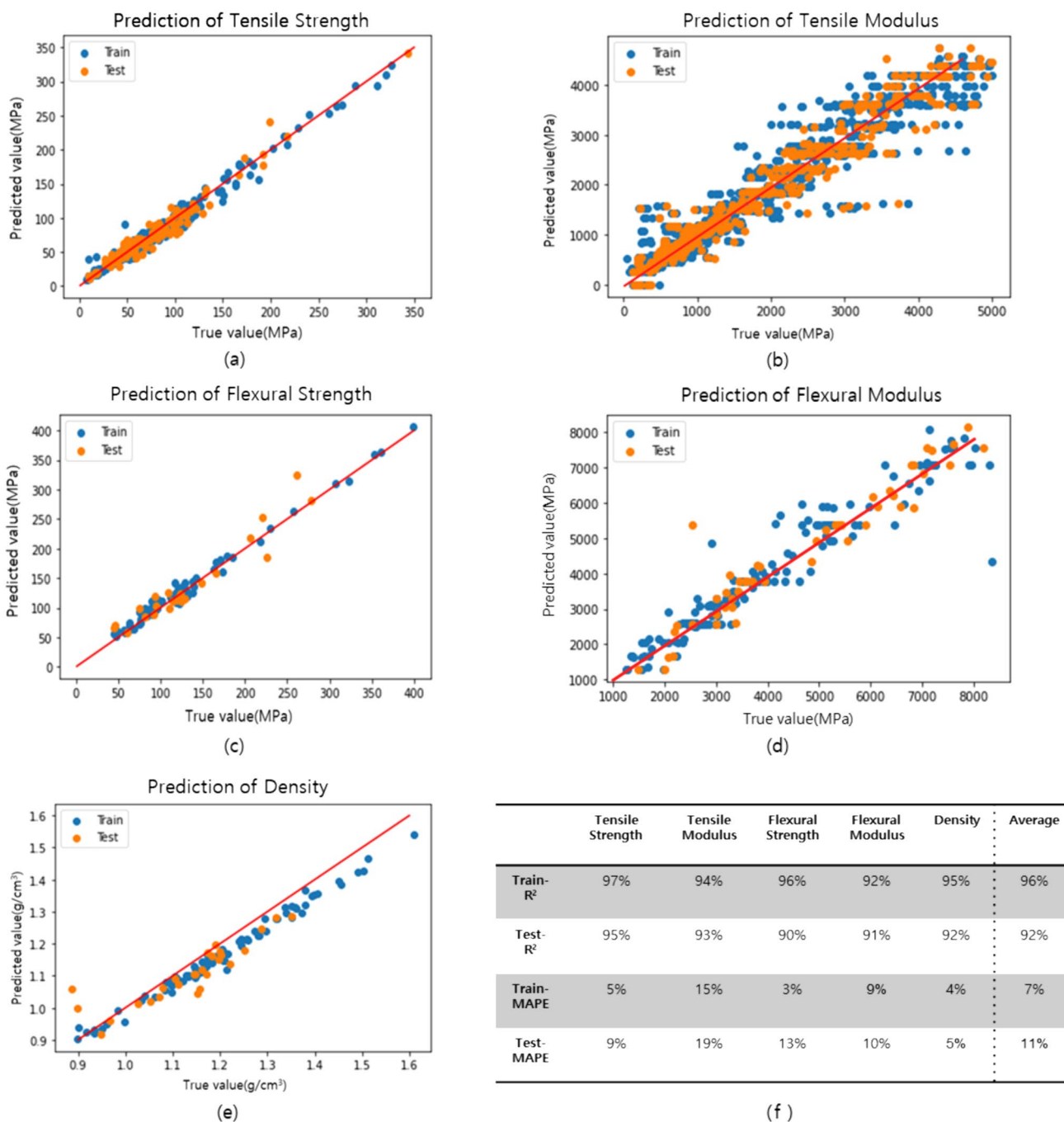
Fig. 10 Mechanical properties prediction results for PMCs

Conclusions

Design of polymer matrix composites depends largely on the experience and intuition of designers, necessitating numerous repetitive experiments and high costs due to the variability in constituent material combinations, processing conditions, and testing environments. This study proposed an integrated mechanical property prediction model

that reflects all the variables such as the combination of constituent materials, processing conditions, and testing conditions to mitigate these issues and effectively utilize accumulated experimental data. Given the limited available data and the need to consider a wide range of input variables, we suggested the use of transfer learning techniques: these techniques involve pre-training on a vast amount of unlabelled data followed by fine-tuning with a smaller set
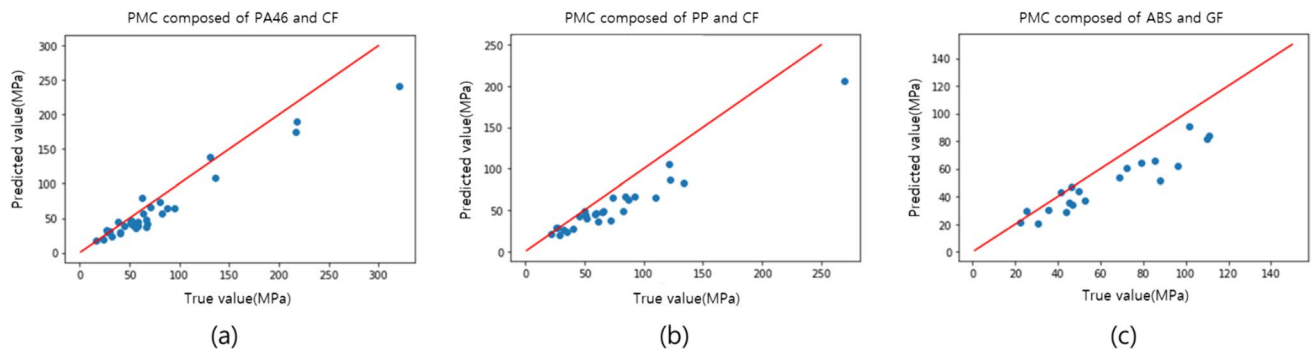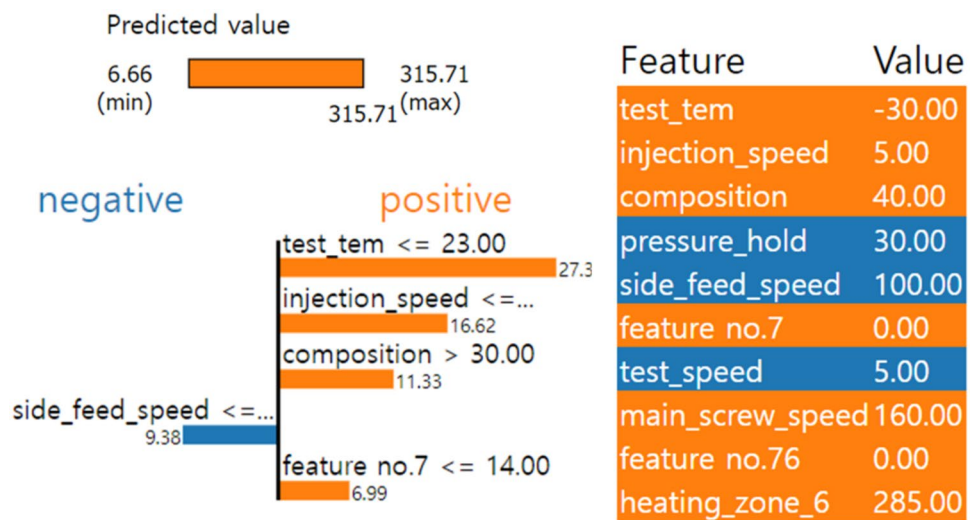
Fig. 11 Tensile strength prediction results analysis for verifying model extrapolation capability. **a** Prediction results for PMC composed of PA46 and CF (R^2: 87%). **b** Prediction results for PMC composed of PP and CF (R^2: 81%). **c** Prediction results for PMC composed of ABS and GF (R^2: 79%)

Fig. 12 XAI analysis results for PMC with highest tensile strength

of labelled data, thereby enhancing the predictive model's performance. The developed model is capable of predicting various mechanical properties like tensile strength, tensile modulus, flexural strength, flexural modulus, and density, demonstrating high accuracy with an average performance of 92% on validation data.

The proposed Transformer-based model was effective in predicting the mechanical properties of polymer composites with limited data, showing a significant performance improvement depending on whether pre-training was involved. It was also verified that in order to enhance the accuracy and reliability of the mechanical property prediction model for polymer composite, it is necessary to utilize training data with variability in composite variables, processing conditions, and testing conditions. The performance of the mechanical property prediction model showed a notable difference of 14% depending on the inclusion or exclusion of processing conditions, underscoring its vital importance.

The model also showed high prediction accuracy for polymer composites not included in the training data, indicating its strong generalization capabilities. Furthermore, The use of LIME in XAI analysis and latent space visualization not only verified that the Transformer model effectively learned the specialized knowledge required for designing polymer composites through data but also confirmed its high suitability as a mechanical property prediction model for polymer composites, effectively acquiring and applying the necessary knowledge. The development of a mechanical property prediction model for polymer composites can be utilized as a predictive model in an automatic inverse design system and is expected to contribute significantly to material innovation across various industrial sectors.

References

1. KPMG, Light weighting of materials: a shift in the center of the automobile light weighting trend. Issue Monit., **96** (2018)

2. M.-Y. Lyu, T.G. Choi, Research trends in polymer materials for use in lightweight vehicles. Int. J. Precis. Eng. Manuf. **16**(1), 213 (2015)

3. H.S. Kim, S.I. Hong, S.J. Kim, On the rule of mixtures for predicting the mechanical properties of composites with homogeneously distributed soft and hard particles. J. Mater. Process. Technol. **112**(1), 109–113 (2001)

4. R. Younes, A. Hallal, F. Fardoun, F. Hajj, Comparative review study on elastic properties modeling for unidirectional composite materials. Compos. Prop. **17**, 391–408 (2012)

5. D. Kaweesa et al., Prediction and validation of composite mechanical properties resulting from voxel-based microstructural design in material jetting. Rapid Prototyp. J. **29**(3), 488–503 (2023)

6. A. Kumar, K. Sharma, A.R. Dixit, A review on the mechanical and thermal properties of graphene and graphene-based polymer nanocomposites: understanding of modelling and MD simulation. Mol. Simul. **46**(2), 136–154 (2020)

7. A. Talapatra, D. Datta, A review of the mechanical, thermal and tribological properties of graphene reinforced polymer nanocomposites: a molecular dynamics simulations methods. Polym. Bull. **80**(3), 2299–2328 (2023)

8. W. Han, J. Zhou, Q. Shi, Research progress on enhancement mechanism and mechanical properties of FRP composites reinforced with graphene and carbon nanotubes. Alex. Eng. J. **64**, 541–579 (2023)

9. J. Liu et al., Machine learning assisted prediction of mechanical properties of graphene/aluminium nanocomposite based on molecular dynamics simulation. Mater. Des. **213**, 110334 (2022)

10. N.X. Ho, T.-T. Le, M.V. Le, Development of artificial intelligence based model for the prediction of Young's modulus of polymer/carbon-nanotubes composites. Mech. Adv. Mater. Struct. **29**(27), 5965–5978 (2022)

11. Z. Zhang, Q. Liu, Wu. Dazhong, Predicting stress–strain curves using transfer learning: knowledge transfer across polymer composites. Mater. Des. **218**, 110700 (2022)

12. A. Gulihonenahali Rajkumar et al., An artificial neural network prediction on physical, mechanical, and thermal characteristics of giant reed fiber reinforced polyethylene terephthalate composite. J. Ind. Textil. **51**(1), 769S-803S (2022)

13. R. Al-Jarrah, F.M. Al-Oqla, A novel integrated BPNN/SNN artificial neural network for predicting the mechanical performance of green fibers for better composite manufacturing. Compos. Struct. **289**, 115475 (2022)

14. M. Iqbal et al., Prediction of residual tensile strength of glass fiber reinforced polymer bars in harsh alkaline concrete environment using fuzzy metaheuristic models. J. Ocean Eng. Sci. **8**(5), 546–558 (2023)

15. R. Hsissou et al., Polymer composite materials: a comprehensive review. Compos. Struct. **262**, 113640 (2021)

16. Cohen, I., Huang, Y., Chen, J., Benesty, J., Benesty, J., Chen, J., et al. Pearson correlation coefficient. Noise Reduct. Speech Process., **1**–4 (2009)

17. P. Sedgwick, Pearson's correlation coefficient. BMJ **345**, e4483 (2012)

18. A. Vaswani et al., Attention is all you need. Adv. Neural. Inf. Process. Syst. **30**, 1 (2017)

19. R. Winter, F. Montanari, F. Noé, D.A. Clevert, Learning continuous and data-driven molecular descriptors by translating equivalent chemical representations. Chem. Sci. **10**, 1692–1701 (2019)

20. Karpov, P., Godin, G., Tetko, I.V., Transformer-CNN: fast and reliable tool for QSAR. arXiv preprint arXiv:1911.06603 (2019)

21. Wang, S., et al., Smiles-bert: large scale unsupervised pre-training for molecular property prediction. In: Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (2019)

22. V. Venkatasubramanian, The promise of artificial intelligence in chemical engineering: Is it here, finally? AIChE J. **65**(2), 466–478 (2019)

23. P. Schwaller et al., Molecular transformer: a model for uncertainty-calibrated chemical reaction prediction. ACS Central Sci. **5**(9), 1572–1583 (2019)

24. Broberg, J., Bånkestad, M., Ylipää, E., Pre-training transformers for molecular property prediction using reaction prediction. arXiv preprint arXiv:2207.02724 (2022)

25. Otsuka, S., et al. PoLyInfo: polymer database for polymeric materials design. In: 2011 International conference on emerging intelligent data and web technologies. IEEE (2011)

26. S. Kim et al., PubChem substance and compound databases. Nucl. Acids Res. **44**(D1), D1202–D1213 (2016)

27. McInnes, L., Healy, J., Melville, J. Umap: uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426 (2018)