

<https://doi.org/10.1038/s41524-025-01532-6>

A physics-enforced neural network to predict polymer melt viscosity

Ayush Jain¹, Rishi Gurnani¹, Arunkumar Rajan¹, H.Jerry Qi² & Rampi Ramprasad¹ ✉

Achieving superior polymeric components through additive manufacturing (AM) relies on precise control of rheology. One rheological property particularly relevant to AM is melt viscosity (η). η is influenced by polymer chemistry, molecular weight (M_w), polydispersity, shear rate ($\dot{\gamma}$), and temperature (T). The relationship of η with M_w , $\dot{\gamma}$, and T is captured by parameterized equations. Several physical experiments are required to fit the parameters, so predicting η of new polymer materials in unexplored physical domains is laborious. Here, we develop a Physics-Enforced Neural Network (PENN) model that predicts the empirical parameters and encodes the parametrized equations to calculate η as a function of polymer chemistry, M_w , polydispersity, $\dot{\gamma}$, and T . We benchmark our PENN against physics-unaware Artificial Neural Network (ANN) and Gaussian Process Regression (GPR) models. We demonstrate that the PENN offers superior values of η when extrapolating to unseen values of M_w , $\dot{\gamma}$, and T for sparsely seen polymers.

Additive Manufacturing (AM) enables the rapid creation of metal or polymer parts with previously unimaginable features and topologies and is therefore poised to disrupt a variety of industries^{1,2}. For polymers, achieving desired properties in the final component is determined by the appropriate choices of material chemistries with suitable rheological properties, as well as conditions adopted during the AM process such as temperature, extrusion rates, etc. At present, a limited palette of chemistries, properties, and conditions is utilized, generally guided by experience, intuition, and empiricism.

In this contribution, we adopt an informatics approach relevant to AM across the chemical and process condition space, to predict one critical rheological property of polymers, namely, the melt viscosity η . Informatics approaches have made major inroads in recent years within materials research^{3–5}, leading to accelerated means for property predictions and providing guidance for the design of new materials^{6–10}. These methods start with available materials data on properties of interest. The materials are then represented numerically to capture and encode their essential features in a machine-readable format. The numerical representations, or fingerprints, are then mapped to available property data using machine learning (ML) algorithms, ultimately producing predictive models for the property considered^{11–15}. Within the AM space, similar methods have been used for final component property prediction¹⁶, process monitoring¹⁷, geometric configuration¹⁸, composition optimization¹⁹, and optimization of printing parameters (albeit mainly for powder-bed AM^{1,18}, but not as much for polymer melt extrusion AM²). Extrusion AM relies on the precise control of polymer melts, which currently requires data from extensive rheological

experiments for each new chemistry. This is a bottleneck in the ink development process². Therefore, predictive capabilities for rheological properties, such as η , are useful to reduce the number of physical experiments aimed at optimization and design.

Melt viscosity of polymers, beyond being a critical property, is attractive to model with ML because there is a reasonable amount of related literature data, although with limited chemical diversity compared to other polymer property datasets^{5,12,15,20,21}. Additionally, there are known physical equations (albeit with empirical parameters) that describe the dependence of η on its governing conditions: temperature (T), average molecular weight (M_w), and shear rate ($\dot{\gamma}$) (Fig. 1A). For instance, it is known that η increases with increasing M_w (via piece-wise power law dependencies), decreases (non-linearly) with increasing $\dot{\gamma}$, and decreases (exponentially) with increasing T . Explicit functional forms and additional background on the behaviors are provided in the Methods section. Molecular weight distributions, quantified by the polydispersity index (PDI), are also known to affect melt viscosity^{22–24}. With this situation in mind, previous works have also addressed the modeling of η using ML^{25–28}. While promising, the majority of these works have focused on specific scenario or are shown to predict unphysical results, making them difficult to apply.

In the present work, we create a physics-enforced neural network (PENN) framework that produces a predictive model of polymer melt viscosity which explicitly encodes the known physical equations while also learning the empirical parameters for new chemistries directly from available data. Physics-informed ML frameworks have shown great promise recently in many application spaces, including atomic modeling, chemistry-

¹School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA. ²School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA. ✉e-mail: rampi.ramprasad@mse.gatech.edu

Enforced Polymer Physics Trends

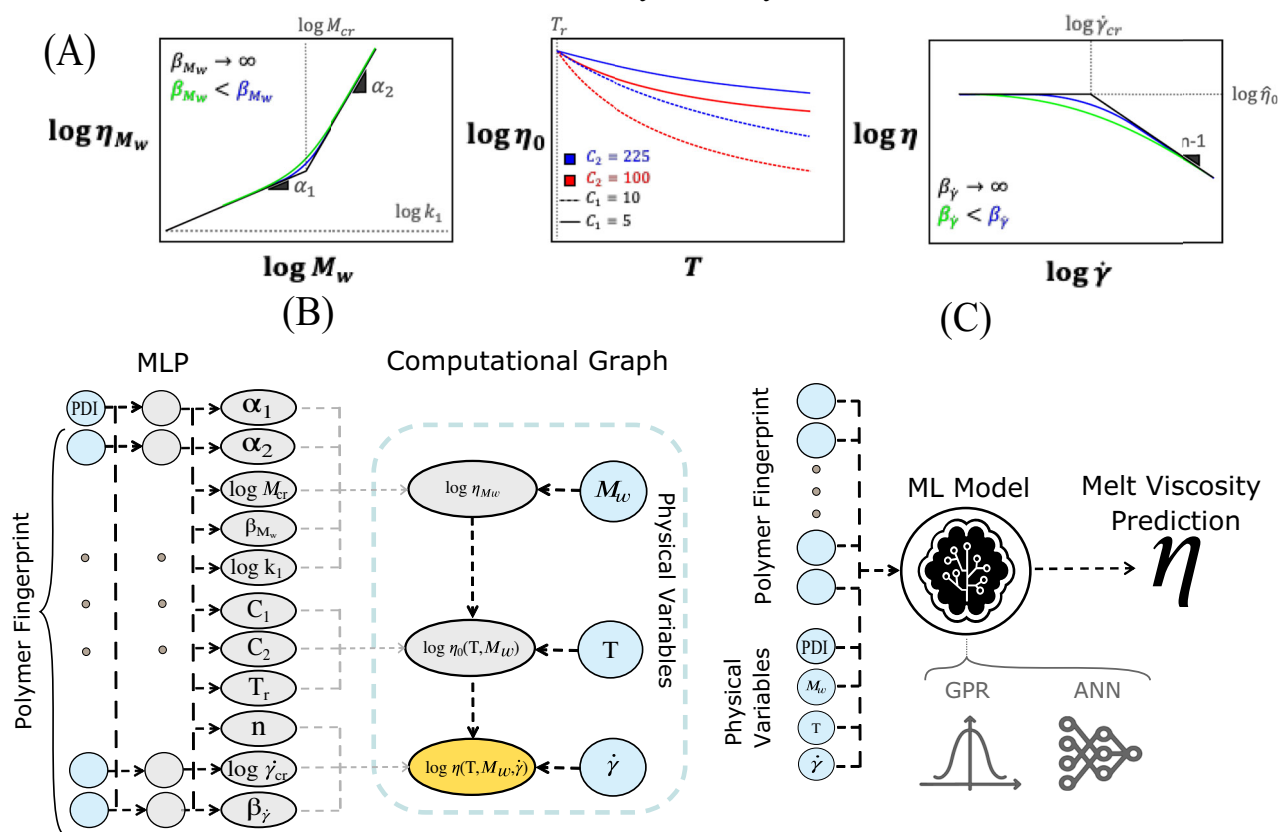


Fig. 1 | The melt viscosity (η) physical trends, learning problems, and machine-learning workflow. **A** Depictions of the functions used to describe the behavior of η with respect to temperature (T), molecular weight (M_w), and shear rate ($\dot{\gamma}$). The functions are parameterized by empirical parameters with physical significance, elaborated in Table 1 and in the Methods section. The η dependence on M_w , is given by $\log \eta_{M_w}$ (Eq. (8) in the Methods section). Empirical parameters define the slopes of the relationship at low M_w (α_1) and high M_w (α_2), the critical molecular weight (M_{cr}), the y -intercept of η_{M_w} (k_1), and the rate of transition from low to high M_w regions (β_{M_w}). The η dependence on T and M_w is given by $\log \eta_0(T, M_w)$ (Eq. (5) in the Methods section), and is parameterized by reference temperature (T_r) and empirical fitting parameters (C_1 and C_2). The effects of C_1 and C_2 are visualized by comparing the trends with different sampled values. The η dependence on $\dot{\gamma}$ is given

by $\log \eta(T, M_w, \dot{\gamma})$ (Eq. (4) in the Methods section). The relevant parameters include shear thinning slope (n), the critical shear rate ($\dot{\gamma}_{cr}$), and the rate of transition from η_0 to shear thinning ($\beta_{\dot{\gamma}}$). **B** The Physics-Enforced Neural Network (PENNN) architecture starts with an input containing the polymer fingerprint and the PDI. A Multi-Layer Perceptron (MLP) uses the concatenated input to predict the empirical parameters. Next, the computational graph uses the predicted empirical parameters to calculate η , via the encoded $\log \eta_{M_w}$, $\log \eta_0(T, M_w)$, and $\log \eta(T, M_w, \dot{\gamma})$ functions. The physical condition variables $\log M_w$, $\log \dot{\gamma}$ and T are input to their respective functions. **C** Physics unaware Artificial Neural Network (ANN) and a Gaussian Process Regression (GPR) are baselines to compare with the PENNN model. The input features to the ANN and GPR models are the concatenated polymer fingerprint, T , M_w , $\dot{\gamma}$, and PDI.

informed materials property prediction, and Physics Informed Neural Networks (PINNs) that solve partial differential equations^{29–33}. Our PENNN for polymer melt viscosity prediction involves a Multi-Layer Perceptron (MLP) that takes as input the polymer chemistry (fingerprinted using our Polymer Genome approach⁶) along with the PDI of the sample, and predicts the empirical parameters as a latent vector (listed in Table 1), used to estimate η as a function of T , M_w , and $\dot{\gamma}$. A computational graph then encodes the dependence of M_w , $\dot{\gamma}$, and T on η (see Fig. 1A) using the equations described in the Methods section. The entire framework (Fig. 1B) is trained on our dataset (elaborated in the Dataset section). The detailed architecture of this framework is described in the Methods section.

We find that this strategy is critical to obtain results that are physically meaningful in extrapolative regimes (e.g., ranges of T , M_w and $\dot{\gamma}$ where there is no training data for chemistries similar to the queried new polymer). This ability is vital given our benchmarking dataset’s sparsity, containing only 93 unique repeat units, although the total number of datapoints is 1903 (including T , M_w , $\dot{\gamma}$, and composition variations). As baselines to assess this PENNN, we trained artificial neural network (ANN) and Gaussian process regression (GPR) models without any physics encoded. We find that the PENNN model is more useful in obtaining credible extrapolative predictions.

Our results indicate that informatics-based data-driven and physics-enforced (when possible) strategies can aid and accelerate extrusion AM innovations in sparse data situations.

Results

Dataset

Melt viscosity data was collected from the PolyInfo repository³⁴ and from the literature cited by PolyInfo. Cited literature data was extracted from tables and figures with the help of the WebPlotDigitizer tool³⁵. The final dataset shown in Fig. 2 includes a total of 1903 datapoints composed of 1326 homopolymer datapoints, 446 co-polymer datapoints, and 113 miscible polymer blend datapoints. The dataset spans a total of 93 unique repeat units with variations in M_w , $\dot{\gamma}$, T , and PDI. For datapoints without a recorded PDI, we impute 2.06, the median PDI of the dataset. Due to inconsistencies with the reporting of polymer structures such as branching and crosslinking for a wide variety of chemistries, the dataset only includes linear polymers.

We found that η at low M_w were underrepresented when compared to η measurements at high M_w . Using the zero-shear viscosity (η_0) relationship with M_w (Fig. 1A), we added 126 datapoints at low M_w (included in the 1903 datapoints). This was achieved by identifying polymer chemistries with

Table 1 | Definitions of empirical parameters predicted by the Physics Enforced Neural Network (PENN)

Parameter	Physical representation	Relevant equation(s)
C_1	Empirical Parameter for the η - T relationship	5
C_2	Empirical Parameter for the η - T relationship	5
T_r	Reference temperature for the η - T	5
M_{cr}	Critical M_w , associated with the onset of polymer chain entanglement	6–8
α_1	Slope of zero-shear viscosity (η_0) vs. M_w when $M_w < M_{cr}$ (approximately 1)	6–8
α_2	Slope of η_0 vs. M_w when $M_w > M_{cr}$ (approximately 3.4)	6–8
β_{M_w}	Measure of transition from α_1 to α_2 at M_{cr}	8
k_1	η_0 when $M = 0$ and $T = T_r$	6–8
$\dot{\gamma}_{cr}$	Critical Shear Rate when $T = T_r$, associated with the onset of shear-thinning	2–4
n	Slope of shear thinning (typically 0.2–0.8 for polymer melts)	2–4
$\beta_{\dot{\gamma}}$	Measure of transition from zero-shear to shear-thinning	4

Their relevance to temperature (T), molecular weight (M_w), and shear rate ($\dot{\gamma}$) when calculating melt viscosity (η) is also provided. The relevant equations in the Methods section are provided for each parameter.

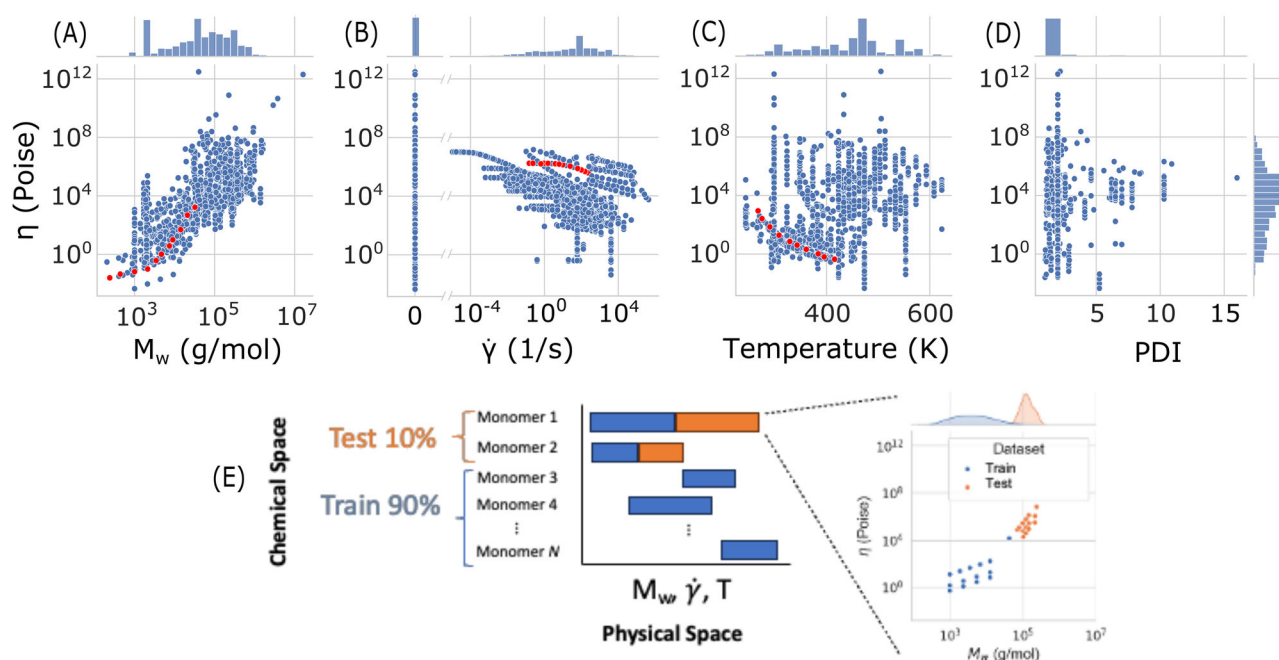


Fig. 2 | Visualizations of the dataset with respect to physical variables. The joint distributions of (A) molecular weight (M_w), (B) shear rate ($\dot{\gamma}$), (C) temperature (T), and (D) polydispersity index (PDI) with respect to melt viscosity (η) are presented. The single distributions for the physical conditions are given on the top axes and the

distribution of η is given on the right-most axis. Each subplot contains all 1903 datapoints from the dataset. (A–C) have highlighted samples in red that exemplify the dependencies depicted in Fig. 1A. E Visual depiction of train-test splitting across chemical space and physical spaces for N monomers in the dataset.

more than five η_0 datapoints at high M_w and a recorded M_{cr} ³⁶. Equation (6) (Methods Section) was fit to each chemistry and extrapolated to estimate η values at low M_w .

Because the viscosity values span several orders of magnitude (Fig. 2), we use the Order of Magnitude Error (OME) to assess ML model accuracy. OME is calculated by taking the Mean Absolute Error of the logarithmically scaled η values. Models with lower OME exhibit more accurate predictions.

Overall assessment of physical intuition with sparse chemical knowledge

An important future use case of our ML models is to estimate the melt viscosity in new physical regimes, given a small amount of knowledge of a given polymer and other chemistries. For example, given a few costly tests of a new polymer at a few molecular weights, one should be able to predict the viscosity at remaining molecular weights, and, likewise, across different

shear rates and temperatures. Figure 2E depicts how this ability was tested through a unique splitting of data into test/train sets across the chemical and physical regimes. First, the monomers were split into train (90%) and test (10%) sets. Within the test monomers, the median of the distributions of the test monomers with respect to a variable in the physical space was calculated. The median was used to split all datapoints containing that monomer: half for a final test split, and the other half for training. The upper or lower half going to testing was randomly chosen. This approach ensures that all the test data focuses on predicting in new physical regimes given a sparse amount of monomer data. This process was repeated three times for each of M_w , $\dot{\gamma}$, and T to ensure that diverse tests were used for evaluation.

Figure 3 shows the combined results of three trials for splits across all three physical variables. Supplementary Fig. 1 shows parity plots that specify the results from each trial. The GPR, ANN, and PENN predictions have acceptable OMEs, indicating that all three can capture some chemical

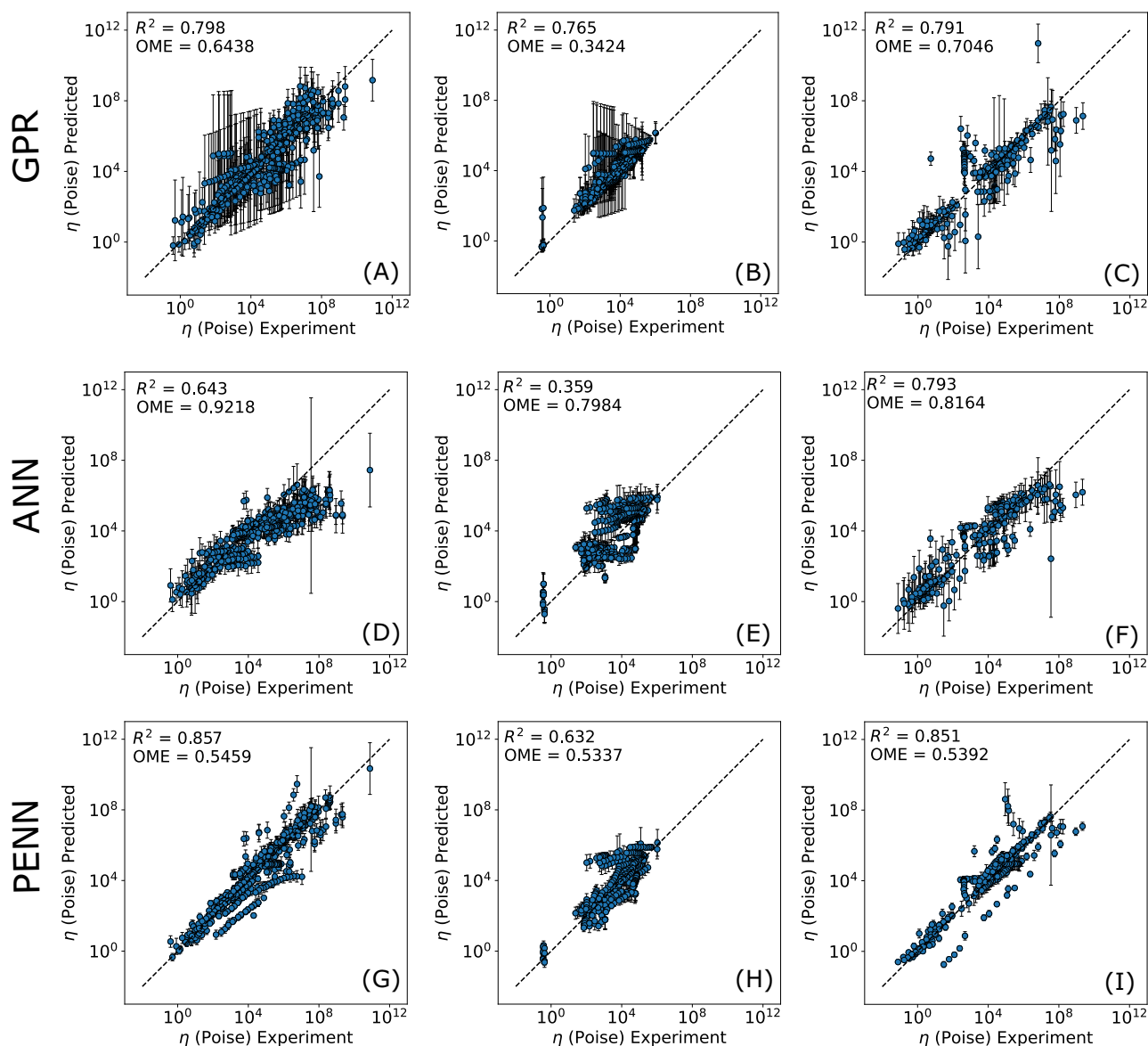


Fig. 3 | Parity plots for unseen viscosity samples. Parity plots are used to assess the models' overall predictive capabilities in new physical regimes based on the physical variable split for molecular weight (M_w), shear rate ($\dot{\gamma}$), and temperature T . Results are compared between Gaussian Process Regression (GPR), Artificial Neural Network (ANN), and Physics Enforced Neural Network (PENN) models. Each plot compares experimental values for melt viscosity (η) to the predicted η across 3

unique test-train splits for each physical variable. The top row (A–C) contains GPR results for (A) the M_w split, (B) the $\dot{\gamma}$ split (C) the T split. The middle row (D–F) contains ANN results for (D) the M_w split, (E) the $\dot{\gamma}$ split (F) the T split. The bottom row (G–I) contains PENN results for (G) the M_w split, (H) the $\dot{\gamma}$ split (I) the T split. The dotted black lines represent perfect predictions. The coefficient of determination (R^2) and Order of Magnitude Error (OME) are reported over these test sets.

information and physical trends. The PENN results in a distinct decrease in OME (an average of 35.97% improvement), and an increase in R^2 (up to 79% for the $\dot{\gamma}$ split) from the ANN. The PENN also outperforms the GPR for the M_w and T splits, but the GPR is more accurate on the test set for $\dot{\gamma}$. In further analysis, we show how the physical viability of these predictions is scrutinized beyond the high-level trends of the parity plot.

Distribution of predicted empirical parameters

Despite the high overall performance of all three models, only the PENN model can produce physically credible predictions in regimes with restricted and sparse data. A comparison of the GPR, ANN, PENN models in estimating crucial empirical parameters (found in Table 1) from sparse data in the held-out set is detailed in Fig. 4.

To establish a benchmark for comparing the three models, we obtained ground truth values of the parameters from the dataset. We did this by identifying subsets of our dataset involving the same polymer with

measured η of several T , M_w , or $\dot{\gamma}$. If a subset contained at least five points, we fitted the corresponding equation (Eqs. (5), (6), or (2)) to obtain empirical parameters. The distributions of these ground truth parameter values are shown in the first row of Fig. 4. There are a limited number of ground truth values because a small number of datapoints satisfy the above conditions. Nevertheless, this small sample set allowed us to make a few inferences about expected viscosity trends. The ground truth values of α_1 and α_2 are close to the theoretical values of 1 and 3.4, respectively³⁷ (background provided in Methods). α_2 values were occasionally less than the expected 3.4, possibly due to outliers or errors in fitting a small number of datapoints. The fitted $\log M_{cr}$ values fell within a range of $10^{2.5} - 10^5$ g/mol. For shear parameters, the majority of samples are found to have n in a range of 0.2 – 0.8, which is typical for polymer melts³⁸. The obtained $\dot{\gamma}_{cr}$ values were found in the range of $10^{-3} - 10^4$ 1/s. The fitted T_r values are mostly in a range of $T_r < 250$ K. This is low when compared to T_g values found in thermal property datasets¹². In our dataset, the datapoints that could be fitted to the η - T relationship were

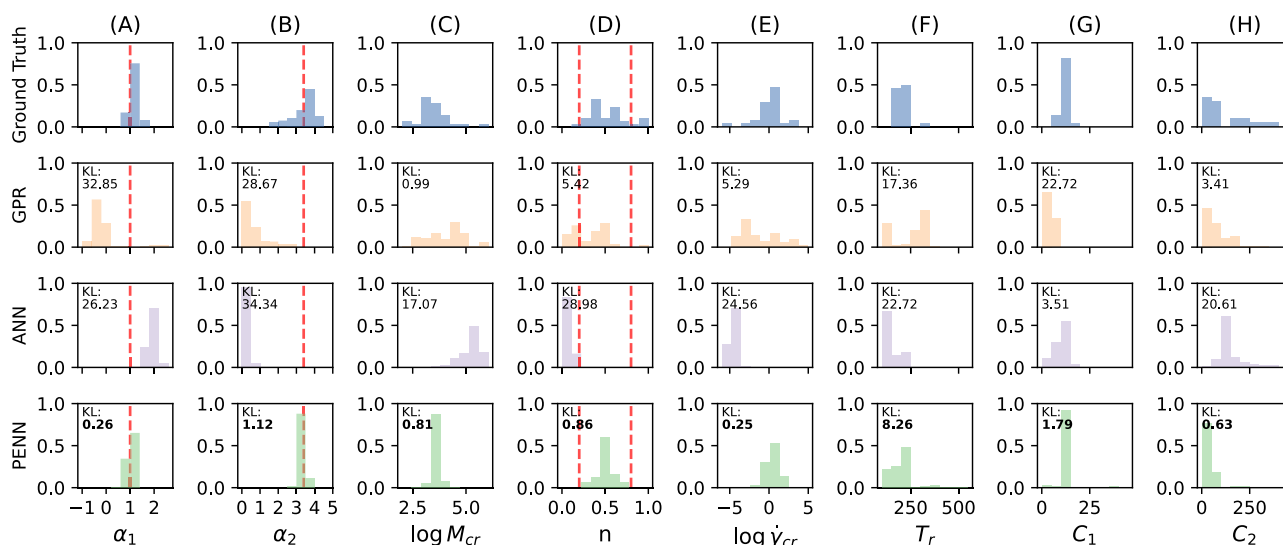


Fig. 4 | Comparisons of empirical parameter predictions of different models. Normalized distributions of empirical parameter values found in the dataset (Ground Truth) are compared to parameter values predicted by Gaussian Process Regression (GPR), Artificial Neural Network (ANN) and Physics Enforced Neural Network (PENN) models. Each column compares a different parameter for the melt viscosity (η) relationship with molecular weight (M_w), shear rate ($\dot{\gamma}$), and temperature (T). The examined parameters include: (A) α_1 , the slope of zero-shear viscosity (η_0) vs. M_w correlation at low M_w (accepted value of 1 depicted by the red dashed line) (B) α_2 , the slope of η_0 vs. M_w at high M_w (accepted value of 3.4 depicted

by the red dashed line), (C) critical molecular weight (M_{cr}), (D) n , the rate of shear thinning (accepted range of 0.2–0.8 depicted by the dashed red lines), (E) critical shear rate ($\dot{\gamma}_{cr}$), (F) reference temperature (T_r) of a polymer. (G, H) show distributions for the C_1 and C_2 fitting parameters for the η - T trend. The ground truth distributions represent 41 samples for M_w parameters, 33 samples for $\dot{\gamma}$ parameters, and 22 samples for T parameters. The Kullback–Leibler (KL) divergence of the model estimation distributions from the ground truth is given in the top left of each histogram. The lowest KL divergence among the three models is bolded for each parameter.

observed at $T < 475\text{K}$, so low T_r values could be overrepresented in the ground truth. The C_1 parameter average was 11.8 and the C_2 parameter average was 159.42 K. This analysis of the ground truth data suggests the desired parameter values our models should predict.

We used two different methods to obtain parameter estimations from the models: one method is unique to the PENN model, and another approach for the purely data-driven ANN and GPR. The PENN model automatically predicts each of the empirical parameters (see Fig. 1B), which are used in the computational graph to predict η . The ANN and GPR do not directly predict the parameters, so we used a fixed extrapolation procedure. The procedure involved selecting an unseen data point and varying a physical variable (one of M_w , $\dot{\gamma}$, and T) within a predetermined range while holding the other two constant. The ranges for each variable encompass similar orders of magnitude as those present in the training dataset (Fig. 2). For M_w extrapolation, a range of $10^2 - 10^7$ g/mol was used to encompass low and high M_w . For shear rate extrapolation, a range of $10^{-5} - 10^6$ 1/s was used to model behaviors in zero-shear and shear-thinning regimes. For temperature extrapolation, ranges of ± 20 K from the original data point's temperature were used to stay within the boundary constraints of Eq. (5). Using this procedure, sets of predictions were made on every unseen datapoint and fit using Eqs. (5), (6), or (2), yielding estimated values of the empirical parameters.

In Fig. 4, we show the feasibility of the models' empirical parameter predictions evaluated against the ground truth values and accepted values (elaborated in the Methods section). For parameters where a theoretical value is well-defined, the Root Mean Square Error (RMSE) of the predictions' deviation from this value is calculated. The parameter prediction distribution is also compared to the ground truth distribution through a discrete Kullback–Leibler (KL) divergence,

$$KL(P \parallel Q) = \sum_i P(i) \log \left(\frac{P(i)}{Q(i)} \right).$$

Intuitively, the KL divergence is a measure of how one probability distribution P deviates from a reference distribution Q over a set of intervals i . A

lower divergence indicates that the predicted parameter distribution is closer to the ground truth. The KL divergence was calculated by finding the entropy between the discretized probability distributions of the ground truth and the ML prediction.

From Fig. 4, it can be seen that GPR struggles to predict expected parameter values. The GPR predictions for α_1 deviate from 1 by an RMSE of 1.26. For some polymers, GPR predicts $\alpha_1 \leq 0$. The GPR predictions for α_2 deviate from 3.4 by an RMSE of 2.87, and are significantly lower than the ground truth values in the dataset. Most predicted values for $\log M_{cr}$ are within the same range as the ground truth, but the proper low and high entanglement behavior is not captured which decreases the credibility of these fittings. For the shear thinning parameter n , some values fall within the expected range of 0.2 – 0.8³⁸ for polymer melts, but others are closer to 0, indicating that the expected shear thinning behavior is not always predicted. The predicted $\dot{\gamma}_{cr}$ distribution is lower than the ground truth, indicating that the GPR forecasts the onset of shear-thinning at a significantly lower $\dot{\gamma}$ than observed (if shear thinning is predicted at all). On temperature dependence, some T_r values are predicted higher than what is seen in the dataset.

The ANN's failure to capture correct physical trends is also evident in the distributions of its fitted parameters. The RMSEs for the ANN's estimated α_1 and α_2 values are 1.31 and 2.79, respectively. ANN overestimates α_1 and underestimates α_2 and therefore does not capture the effects of high M_w chain entanglement. The ANN predictions estimate a low n for a subset of polymers, which goes against the definition of shear thinning. The predicted $\dot{\gamma}_{cr}$ values are lower than the ground truth distribution, indicating that the ANN struggles to capture the shear-thinning transition region from the dataset. The ANN predictions for the T trend are closest to the ground truth in comparison to its trends of the other variables, because T is a smoother, exponential function (Fig. 1A), enabling an easier average fitting.

The PENN outperforms the ANN in estimating feasible empirical parameters as depicted by lower KL Divergence values in the last row of Fig. 4. The RMSEs of the predicted α_1 and α_2 values are 0.05 and 0.17, which are substantially smaller than that of the ANN. Moreover, all the predicted values of $\log M_{cr}$ are within the ground truth range of 2.5 – 5. The PENN model can also learn the correct shear thinning phenomenon by predicting

n values between 0.2 – 0.8³⁸ and a $\dot{\gamma}_{cr}$ distribution that mirrors the dataset. The PENN's predicted range of T , is closest to the ground truth. For the C_1 parameter, the PENN predicted distribution is closest to the proposed value of $C_1 = 7.60$ (detailed in the Methods section), also having the lowest divergence from the ground truth. For C_2 predictions, although the KL Divergence of the PENN is lower than the ANN, the PENN is confined to much lower values of C_2 , and has an average much lower than some experimentally derived values, such as $C_2 = 227.3$ K³⁹.

Overall, the average KL divergence across all parameter distributions for the GPR, ANN, and PENN are 14.59, 22.24, and 1.74, respectively. The overall distributions of empirical parameters points to the PENN having

greater capabilities for producing physically correct results, than a purely data-driven model.

Performance in extrapolative regimes

In Table 2, we summarized the performance of predicted η profiles over wide ranges of M_w (256 extrapolations), $\dot{\gamma}$ (71 extrapolations), and T (127 extrapolations) for all three models considered. We define a successful extrapolation as a model that is able to predict the correct trends while maintaining accuracy over the train and test points. In this study, a prediction is considered accurate for an experimental point if the experimental data falls reasonably within the predicted uncertainty bounds. However, in practice the required precision may vary depending on the specific practical use case. Overall, the PENN successfully predicts 80.4% of M_w extrapolations, 49.2% of $\dot{\gamma}$ extrapolations, and 54.1% of T extrapolations. The ANN rarely achieves correct physical trends for M_w , or $\dot{\gamma}$ extrapolations in the span of the dataset and only predicts successful profiles for 17.2% of T extrapolations. The GPR model also exhibits a low performance in extrapolation. There are several instances (given in the brackets in Table 2) where the ANN and GPR successfully fit the data points but fail to extrapolate correctly beyond the dataset. This underscores the need for information beyond experimental data to enable extrapolation to new physical regimes.

Figure 5 shows a few examples of the extrapolation results summarized in Table 2. A much larger set of examples of both successful and unsuccessful extrapolations by the PENN compared to the GPR and the ANN are provided in Supplementary Figs. 2–7. Figure 5A–C shows examples of the

Table 2 | Extrapolative predictive performance of the PENN, ANN, and GPR models along the unseen molecular weight (M_w), shear rate ($\dot{\gamma}$), and temperature T regimes

Model	M_w	$\dot{\gamma}$	T
PENN	80.4%	49.2%	54.1%
ANN	4.30% [64.4%]	4.22% [19.7%]	17.2% [19.5%]
GPR	0.0% [68.0%]	7.04% [40.8%]	7.03% [28.1%]

The values within brackets for the ANN and GPR show the percentages of extrapolations where the unseen data was predicted correctly, but the extrapolated trend beyond available data regimes was physically incorrect.

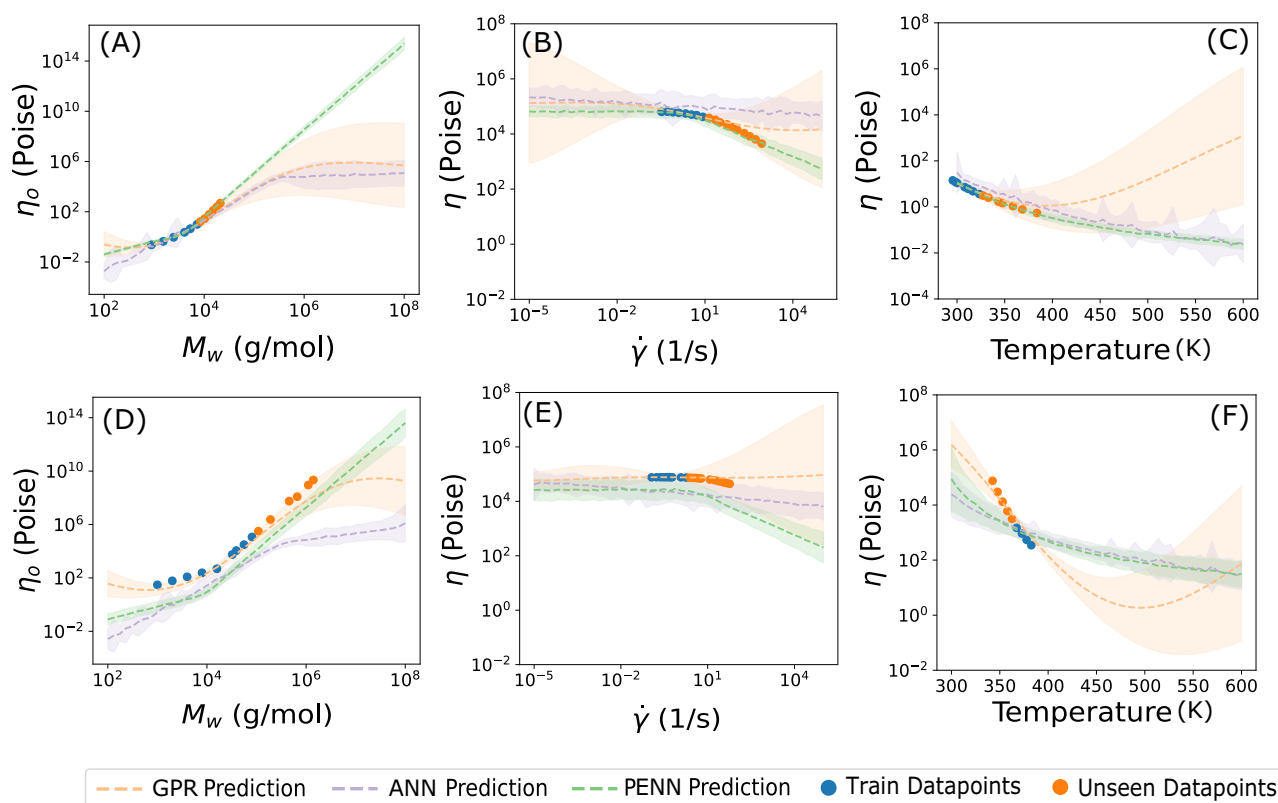


Fig. 5 | Example extrapolations of melt viscosity for different physical variables. Examples of successful (A–C) and unsuccessful (D–F) melt viscosity (η) and zero-shear melt viscosity (η_0) predictions over wide ranges of molecular weight (M_w), shear rate ($\dot{\gamma}$), temperature (T) by the Physics Enforced Neural Network (PENN) models. The extrapolated predictions are compared to those by Gaussian Process Regression (GPR) and Artificial Neural Network (ANN) models given the same training information. **A** is a good η_0 - M_w extrapolation for [*]CCCCCCCCCOC(=O)CCCC(=O)O[*] at $T = 382.15$ K. **B** is a good η - $\dot{\gamma}$ extrapolation for a copolymer of [*]CC[*]CC(C)C and [*]CC[*]CCCCCCC (0.968:0.032) ($M_w = 290000$ g/mol, PDI = 7.8) at $T = 543.15$ K. **C** is a good η - T extrapolation for [*]CCOCCOC-COC(=O)CCCCCCCC(=O)O[*] ($M_w = 2000$ g/mol, $\dot{\gamma} = 60$ 1/s). **D** is an unsuccessful η_0 - M_w extrapolation for [*]C=CCC[*] at $T = 490.15$ K, with possible mispredictions of M_{cr} and k_1 . **E** is an unsuccessful η - $\dot{\gamma}$ extrapolation for a copolymer of [*]C[*] and [*]CC[*]OC(C) (0.72:0.28) ($M_w = 60000$ g/mol), with possible misprediction of $\dot{\gamma}_{cr}$ and η_0 . **F** is an unsuccessful η - T extrapolation for [*]CC(O)COc1ccc(C(C)(C)c2ccc(O[*]cc2)cc1 ($M_w = 1696$ g/mol, $\dot{\gamma} = 0.0$ 1/s) with a possible misprediction of T .

PDI = 7.8) at $T = 543.15$ K. **C** is a good η - T extrapolation for [*]CCOCCOC-COC(=O)CCCCCCCC(=O)O[*] ($M_w = 2000$ g/mol, $\dot{\gamma} = 60$ 1/s). **D** is an unsuccessful η_0 - M_w extrapolation for [*]C=CCC[*] at $T = 490.15$ K, with possible mispredictions of M_{cr} and k_1 . **E** is an unsuccessful η - $\dot{\gamma}$ extrapolation for a copolymer of [*]C[*] and [*]CC[*]OC(C) (0.72:0.28) ($M_w = 60000$ g/mol), with possible misprediction of $\dot{\gamma}_{cr}$ and η_0 . **F** is an unsuccessful η - T extrapolation for [*]CC(O)COc1ccc(C(C)(C)c2ccc(O[*]cc2)cc1 ($M_w = 1696$ g/mol, $\dot{\gamma} = 0.0$ 1/s) with a possible misprediction of T .

PENN correctly extrapolating η given a small amount of information about a monomer in another part of the physical regime in unseen regimes. The ANN and GPR models are uncertain in these unseen regimes, resulting in large confidence intervals. In Fig. 5A, the PENN model accurately predicts the region near M_{cr} where the η - M_w relationship transitions from unentangled to entangled, and can therefore accurately predict η values at high M_w , despite not having seen any data in this region. The errors for the ANN and GPR in Fig. 5A are low, within approximately an order of magnitude of error. However, the ANN predictions have a near-constant slope around M_{cr} (implying $\alpha_1 \approx \alpha_2$) and are inconsistent with the effects of polymer chain entanglements at high M_w . The GPR model also fails to predict a higher α_2 slope. In Fig. 5B, only the PENN model predicts a zero-shear and shear-thinning region when predicting the η - $\dot{\gamma}$ relationship of the given copolymer. The GPR model fits the training points but mispredicted shear-thinning at high shear rates. The ANN model predicts a decreasing relationship consistent with shear-thinning but doesn't predict the zero-shear region. This could be an example of spectral bias within neural networks, which describes how ANNs prioritize global or "low frequency" patterns in data over local or "high frequency" patterns⁴⁰. The general decreasing trend of η - $\dot{\gamma}$ is "low frequency" and is captured by the ANN. In contrast, the transition regions are of a "higher frequency" and are not captured by the ANN. In Fig. 5C, the PENN model predicts the correct η - T relationship. The ANN model also predicts an exponential relationship but with a higher inaccuracy. The GPR model fits both the training and unseen datapoints, but predicts an unphysical trend beyond this. Overall, the PENN model makes predictions that follow the expected behaviors (Fig. 1A) of polymer melts.

Correctly extrapolated samples by the PENN model, such as the ones in Fig. 5A–C make up 67.5% of the extrapolated test cases, which is a significant improvement relative to both the ANN and GPR. The PENN model also has room for improvement, especially when applied to datasets with low chemical diversity. Overfitting to a small set of chemistries in training can lead to the inaccurate prediction of parameters when making predictions for unseen chemistries. This behavior is demonstrated in Fig. 5D–F, where the PENN predicts a plausible rheological trend but incorrect values for unseen polymers. However, the PENN model introduces a layer of interpretability unavailable to physics-unaware models. Based on the predictions we can reasonably infer which parameters were over- or under-estimated. In Fig. 5D, the PENN model predicts near-correct α_1 and α_2 slopes, but the predicted M_{cr} and k_1 values are underestimated. Figure 5E depicts how an underestimated η_0 (caused by inaccuracies in predicted M_{cr} , α_1 , α_2 and/or k_1) can cause inaccurate η predictions for all other $\dot{\gamma}$ values. We also see this phenomenon in Fig. 5F, where T_c is likely underestimated. The propagating error causes the PENN model to predict an inaccurate trend across the entire spectrum of T . Despite these errors, the pinpointing of the PENN's weak spots can be used to add targeted training data to improve the model. This level of interpretation is unique to the PENN and cannot be done for the GPR and ANN.

These examples of extrapolations provide insights into the applicability of PENN versus pure data-driven methods when using datasets that contain limited chemistries. The equations used in the PENN are based on assumptions and generalizations, and may not account for all physical nuances. These must be considered when applying PENNs to future material design and process optimization problems.

Discussion

In this study, we introduce a Physics Enforced Neural Network (PENN), a strategy that combines data-driven techniques with established empirical equations, to predict the melt viscosity of polymer melts with better physics-guided generalization and extrapolation. The PENN makes predictions across many chemical compositions and relevant physical parameters, including molecular weight, shear rate, temperature, and polydispersity index. We compared our PENN approach against the purely data-driven, physics-unaware, Artificial Neural Network and Gaussian Process

Regression. In extrapolative regimes, our PENN model outperforms the physics-unaware counterparts and offers an elevated level of interpretability and generalizability. To enhance generalizability across chemistries, future work could increase the chemical space in the dataset through new experiments, molecular dynamics simulations, and/or more aggressive data acquisition from literature.

This work has profound implications for additive manufacturing (AM) and materials informatics. The PENN model's capability to guide the rheological control of diverse polymer resins accelerates the development of new printing materials, thereby expanding AM's utility. Our methodology offers a blueprint for modeling other properties governed by empirical equations. The initial success of the PENN architecture for melt viscosity is a powerful demonstration of how data-driven insights combined with established knowledge can propel us into a new era of rapid advancements in materials science and engineering.

Methods

Fingerprinting and feature engineering

The chemical attributes of a polymer are represented by a unique fingerprinting scheme. The fingerprints (FPs) contain features derived from atomic-level, block-level, chain-level, and morphological descriptors of a polymer as described at length earlier⁶. The dataset contains homo- and co-polymers, and miscible polymer blends. Co-polymers and blends contain multiple repeating units, each with a separate FP. For co-polymers, the FP of each unit was aggregated to a single copolymer FP using a weighted average (with weight equal to composition percentage)¹². Similar to previous work¹², all co-polymers were treated as random. For miscible polymer blends, the FP of each unit was aggregated to a single FP using a weighted harmonic average (with the weight equal to composition percentage)²⁰. For blends containing units with different M_w and/or PDI, the weighted average over each unit was used.

Enforced polymer physics trends

In this section, we detail the physics-based correlations included within the Physics Enforced Neural Network (PENN).

We enforce dependencies of η on temperature (T), molecular weight (M_w), and shear rate ($\dot{\gamma}$) through $\eta(M_w, T, \dot{\gamma})$, which we derive below.

Preamble: smoothing of piecewise functions. When going from one function $g(a, b)$ in a low regime ($a < b$) to another function $h(a, b)$ in a high regime ($a > b$), we can use the smoothed Heaviside step function,

$$H_\beta = \frac{1}{1 + \exp(-\beta x)}, \quad (1)$$

where β is a tunable rate of transition.

A function $f(a, b)$ that transitions from $g(a, b)$ to $h(a, b)$ is given by

$$f(a, b) = g(a, b) \times H_\beta(b - a) + h(a, b) \times H_\beta(a - b)$$

η dependence on $\dot{\gamma}$, T , and M_w . The η dependence on $\dot{\gamma}$ follows the physics of shear-thinning fluids^{41–43}. In these fluids, at low $\dot{\gamma}$, there is not enough force between chains to break entanglements and cause movement, so η remains constant at η_0 . At a critical shear rate, $\dot{\gamma}_{cr}$, the shear force is high enough to cause chain alignment, making chain diffusion easier. Beyond $\dot{\gamma}_{cr}$, η decreases according to a shear-thinning linear power law⁴¹. This trend can be represented by a function (Equation (2)) across both the zero-shear and shear thinning regimes^{41,44–46},

$$\eta(M_w, T, \dot{\gamma}) = \frac{\eta_0(M_w, T)}{\left(1 + \frac{\dot{\gamma}}{\dot{\gamma}_{cr}}\right)^{1-n}} \quad (2)$$

$$\log \eta = \log \eta_0(M_w, T) + (n - 1) \log \left(1 + \frac{\dot{\gamma}}{\dot{\gamma}_{cr}}\right)$$

where the parameter n describes the sensitivity to shearing⁴⁷. For shear-thinning fluids, $n < 1$. For most polymer melts, n is empirically known to be in the range 0.2–0.8³⁸.

Equation (2) is unfavorable to use directly because $\dot{\gamma}$ spans several orders of magnitude, so $\log \dot{\gamma}$ must be used as an input. Eq. (2) cannot be adapted to use $\log \dot{\gamma}$ as an input (due to the +1 in the denominator), so we depict the relationship across the low $\dot{\gamma}$ and high $\dot{\gamma}$ regimes as a piecewise function on the log-scale,

$$\log \eta = \begin{cases} \log \eta_0 & \text{if } \dot{\gamma} \ll \dot{\gamma}_{cr} \\ \log \eta_0 + (n - 1) \log\left(\frac{\dot{\gamma}}{\dot{\gamma}_{cr}}\right) & \text{if } \dot{\gamma} \gg \dot{\gamma}_{cr} \end{cases} \quad (3)$$

We smooth Eq. (3) with $H_{\beta_{\dot{\gamma}}}$ to get $\log \eta(M_w, T, \dot{\gamma})$ (Eq. (4)),

$$\log \eta(M_w, T, \dot{\gamma}) = \log \eta_0(M_w, T) \times H_{\beta_{\dot{\gamma}}}(\log \dot{\gamma}_{cr} - \log \dot{\gamma}) + \left(\log \eta_0(M_w, T) + (n - 1) \log\left(\frac{\dot{\gamma}}{\dot{\gamma}_{cr}}\right) \right) \times H_{\beta_{\dot{\gamma}}}(\log \dot{\gamma} - \log \dot{\gamma}_{cr}), \quad (4)$$

where $\beta_{\dot{\gamma}}$ is a parameter that dictates the rate of shift from zero-shear to shear-thinning. For our implementation, we found that optimization over the $\dot{\gamma}$ domain was optimal when $\beta_{\dot{\gamma}} = 30$.

$\log \eta_0(M_w, T)$ is defined by the T dependence. As temperature increases, so does the rate of molecular self-diffusion, resulting in lower η seen in fluidic polymer melts⁴³. The William–Landel–Ferry (WLF) equation^{39,48} describes the exponential decrease in η as the temperature increases. Therefore, we can encode temperature dependence as

$$\eta_0 = \eta_{M_w} \times 10^{\frac{-C_1(T-T_r)}{C_2+(T-T_r)}}, \forall T \geq T_r \quad (5)$$

$$\log \eta_0(M_w, T) = \log \eta_{M_w} + \frac{-C_1(T-T_r)}{C_2+(T-T_r)}, \forall T \geq T_r$$

where T_r is a reference temperature and C_1 and C_2 are material-dependent empirical parameters. The values for these are dependent on polymer chemistry. $C_1 = 7.60$ and $C_2 = 227.3$ K are examples of values that have been proposed³⁹ from observations of experiments on a small subset of polymers. The reference temperature T_r is within a few degrees of the glass transition temperature T_g . It has been proposed that the WLF relationship holds within the range of T_g to $T_g + 200\text{K}$ ³⁹.

η_{M_w} is defined by the M_w dependence. Longer and heavier polymer chains experience increased entanglements, which hinder chain reptation in the polymer melt at low shear^{37,43}. Equation (6) is a piece-wise power law that describes this phenomenon.

$$\eta_{M_w} = \begin{cases} k_1 M_w^{\alpha_1} & \text{if } M_w < M_{cr} \\ k_2 M_w^{\alpha_2} & \text{if } M_w \geq M_{cr} \end{cases} \quad (6)$$

where

$$k_2 = k_1 M_{cr}^{\alpha_1 - \alpha_2}.$$

M_{cr} is the critical molecular weight, above which entanglement density is high enough to increase the impact of M_w on η_0 . The two power laws intersect at $M_w = M_{cr}$ ³⁷. M_{cr} is found to be approximately 2–4 times the molecular weight at which chain entanglement starts, but the exact value is polymer dependent⁴³. α_1 is the slope of the $\log \eta_0$ - $\log M_w$ curve if $M_w < M_{cr}$ and α_2 is the slope if $M_w \geq M_{cr}$. Typically, α_1 is theoretically and empirically determined to be about 1, while α_2 is found to be about 3.4^{37,43}, but the exact value is dependent on the polymer. k_1 and k_2 are the y -intercepts of each power law and are polymer-dependent.

M_w and η_0 span several orders of magnitude, so we use Eq. (6) in the log-scale to get Eq. (7),

$$\log \eta_{M_w} = \begin{cases} \log k_1 + \alpha_1 \log M_w & \text{if } M_w < M_{cr} \\ \log k_1 + (\alpha_1 - \alpha_2) \log M_{cr} + \alpha_2 \log M_w & \text{if } M_w \geq M_{cr} \end{cases} \quad (7)$$

Smoothing Eq. (7) with $H_{\beta_{M_w}}$ gives Eq. (8),

$$\log \eta_{M_w} = [\log k_1 + \alpha_1 \log M_w] * H_{\beta_{M_w}}(\log M_{cr} - \log M_w) + [\log k_1 + (\alpha_1 - \alpha_2) \log M_{cr} + \alpha_2 \log M_w] * H_{\beta_{M_w}}(\log M_w - \log M_{cr}), \quad (8)$$

where $H_{\beta_{M_w}}$ is the smoothed Heaviside step function using β_{M_w} , a parameter which dictates the rate of shift from α_1 to α_2 .

Therefore, Eqs. (4), (5), and (8) determine the $\log \eta(M_w, T, \dot{\gamma})$. The predicted parameters n , $\dot{\gamma}_{cr}$, $\beta_{\dot{\gamma}}$ determine the $\dot{\gamma}$ dependence in $\log \eta(M_w, T, \dot{\gamma})$, which is also a function of $\eta_0(M_w, T)$. The predicted parameters C_1 , C_2 , and T_r determine the T dependence in $\eta_0(M_w, T)$, which is also a function of η_{M_w} . The predicted parameters α_1 , α_2 , M_{cr} , β_{M_w} , and k_1 determine the M_w dependence in η_{M_w} . The parameter outputs of the MLP have physically appropriate bounding ranges (reported in Supplementary Table 1).

η dependence on PDI. The dispersity of molecular weights in a polymer melt affects the bulk motion of polymer chains⁴³. For example, a short and long chain may diffuse differently compared to two medium-sized chains. Therefore, using just the M_w without any knowledge of dispersity can mislead the ML model. We account for dispersity by using the polydispersity index (PDI),

$$PDI = \frac{M_w}{M_n},$$

where M_n is the number average molecular weight. Empirical models for this relationship^{22,24} may require detailed information on the specific shape of the molecular weight distribution of a polymer melt. Not all of our data points contain proper information on PDI (as discussed in the Results section), so we do not directly encode η -PDI trends within the computational graph. Instead, the PDI could affect the transitions in the critical regimes of the η_0 - M_w relationship and the η - $\dot{\gamma}$ relationship (when $\dot{\gamma} = \dot{\gamma}_{cr}$ or $M_w = M_{cr}$)^{22–24}. We incorporate this effect through the parameters β_{M_w} and $\beta_{\dot{\gamma}}$ (described in Table 1). A higher value of β_{M_w} or $\beta_{\dot{\gamma}}$ creates a quicker transition within their respective critical regimes.

PENN training

This entire PENN architecture is trained, in part, to minimize the error of viscosity predictions. The sum of these errors across all n training points is called the viscosity loss \mathcal{L}_η , defined in Eq. (9). Each data point is denoted by its index i .

$$\mathcal{L}_\eta = \frac{1}{n} \sum_{i=1}^n (\hat{\eta}_i - \eta_i)^2 \quad (9)$$

During training, we add loss terms (see Eq. (10)) to penalize the predicted α_1 and α_2 for the i th training point ($\hat{\alpha}_{1,i}$ and $\hat{\alpha}_{2,i}$, respectively) for deviating from their average values. The viscosity loss plus the penalty terms form the total loss \mathcal{L} .

$$\mathcal{L} = \mathcal{L}_\eta + \frac{1}{n} \sum_{i=1}^n w_\alpha [(\hat{\alpha}_{1,i} - 1)^2 + (\hat{\alpha}_{2,i} - 3.4)^2] \quad (10)$$

w_α is a hyperparameter that controls the impact that known values of the α_1 and α_2 parameters have on the final loss.

Machine learning approaches

The PENN and ANN models were implemented in PyTorch⁴⁹. All models were trained on the same 9:1 (Train:Test) split. Before training, the features and η were scaled to a range of $(-1,1)$. The polymer fingerprint, PDI, and temperature were scaled with the Scikit-Learn MinMaxScaler⁵⁰ to a range of $(-1,1)$. The $\dot{\gamma}$ was scaled by first adding a small value of 10^{-5} , taking the \log_{10} , and then scaling to $(-1,1)$. M_w was scaled by taking the \log_{10} value and then scaling to $(-1,1)$. For the PENN, $\log M_w$ and $\log \dot{\gamma}$ use the same scaling bounds as η .

Within the training set, a 10-fold cross-validation (CV) was used to ensure that the models did not overfit the training set. The ANN and PENN models also had separate models trained for each CV split. Hyperparameter optimization was performed using the Hyperband⁵¹ optimization algorithm over each CV fold for both the ANN and the PENN models, with RayTune⁵² implementations, respectively. The ANN and PENN models, both containing 4 layers (including 2 hidden layers), involved optimization of the same hyperparameters: layer 1 size (64, 128, 256, 512), layer 1 dropout (0,0.01, 0.015,0.02,0.025,0.03), layer 2 size (64, 128, 256, 512), layer 2 dropout (0,0.01,0.015,0.02,0.025,0.03), and weight decay (0.00001, 0.00005, 0.0001, 0.0005, 0.001). For the PENN, w_α (0.001, 0.005, 0.01, 0.03, 0.05) was also optimized. The value corresponding to the lowest \mathcal{L}_η (Eq. (9)) of the CV test split was used.

The Adam optimizer was used to train the models with a learning rate (LR) reduction by a factor or 0.5 on the plateau of the validation loss given a patience of 20 epochs. An initial LR of 0.0001 was used for the PENN. Empirically, we found that the PENN tuning was sensitive to high LR. The initial LR for the ANN was 0.001. Training was stopped with an Early Stopping patience of no improvement in the validation loss after 25 epochs.

The GPR model was implemented using Scikitlearn⁵⁰ trained using Bayesian optimization to tune key hyperparameters. The hyperparameters optimized include the noise level (`alpha`) with a range of $[10^{-2}, 10^1]$, the length scale of the RBF kernel (`length_scale`) with a range of $[10^{-2}, 10^2]$, and the constant value used in the kernel (`constant_value`) with a range of $[10^{-2}, 10^2]$, each with a logarithmic uniform prior. The optimization was performed over 50 iterations each over the 10-fold cross-validation, with the best-performing model parameters selected based on the results. The scaling for the inputs and outputs of the GPR were the same as the ANN.

Data availability

The dataset used in this study is available on the Ramprasad group's github (https://github.com/Ramprasad-Group/polyVERSE/tree/main/Other/Melt_Viscosity).

Code availability

A sample framework for the PENN architecture is available at https://github.com/Ramprasad-Group/PENN_Melt_Viscosity. Python 3.10 and PyTorch 1.12.1 were used.

Received: 8 October 2024; Accepted: 10 January 2025;

Published online: 20 February 2025

References

- Wang, C., Tan, X., Tor, S. & Lim, C. Machine learning in additive manufacturing: state-of-the-art and perspectives. *Addit. Manuf.* **36**, 101538 (2020).
- Das, A., Gilmer, E. L., Biria, S. & Bortner, M. J. Importance of polymer rheology on material extrusion additive manufacturing: correlating process physics to print properties. *ACS Appl. Polym. Mater.* **3**, 1218–1249 (2021).
- Chen, L. et al. Polymer informatics: current status and critical next steps. *Mater. Sci. Eng.: R: Rep.* **144**, 100595 (2021).
- Batra, R., Song, L. & Ramprasad, R. Emerging materials intelligence ecosystems propelled by machine learning. *Nat. Rev. Mater.* **6**, 655–678 (2021).
- Tran, H. et al. Design of functional and sustainable polymers assisted by artificial intelligence. *Nat. Rev. Mater.* **9**, 866–886 (2024).
- Doan Tran, H. et al. Machine-learning predictions of polymer properties with Polymer Genome. *J. Appl. Phys.* **128**, 171104 (2020).
- Mannodi-Kanakthodi, A. et al. Scoping the polymer genome: a roadmap for rational polymer dielectrics design and beyond. *Mater. Today* **21**, 785–796 (2018).
- Chen, L., Kern, J., Lightstone, J. P. & Ramprasad, R. Data-assisted polymer retrosynthesis planning. *Appl. Phys. Rev.* **8**, 031405 (2021).
- Gurnani, R. et al. PolyG2G: A novel machine learning algorithm applied to the generative design of polymer dielectrics. *Chem. Mater.* **33**, 7008–7016 (2021).
- Liu, D.-F., Feng, Q.-K., Zhang, Y.-X., Zhong, S.-L. & Dang, Z.-M. Prediction of high-temperature polymer dielectrics using a Bayesian molecular design model. *J. Appl. Phys.* **132** (2022).
- Lorenzini, R., Kline, W., Wang, C., Ramprasad, R. & Sotzing, G. The rational design of polyurea I& polyurethane dielectric materials. *Polymer* **54**, 3529–3533 (2013).
- Kuenneth, C., Schertzer, W. & Ramprasad, R. Copolymer informatics with multitask deep neural networks. *Macromolecules* **54**, 5957–5961 (2021).
- Moriwaki, H., Tian, Y.-S., Kawashita, N. & Takagi, T. Mordred: a molecular descriptor calculator. *J. Cheminformatics* **10**, 1–14 (2018).
- Gurnani, R., Kuenneth, C., Toland, A. & Ramprasad, R. Polymer informatics at scale with multitask graph neural networks. *Chem. Mater.* **35**, 1560–1567 (2023).
- Gurnani, R. et al. AI-assisted discovery of high-temperature dielectrics for energy storage. *Nat. Commun.* **15**, 6107 (2024).
- Bayraktar, Ö., Uzun, G., Çakiroğlu, R. & Guldaz, A. Experimental study on the 3D-printed plastic parts and predicting the mechanical properties using artificial neural networks. *Polym. Adv. Technol.* **28**, 1044–1051 (2017).
- Delli, U. & Chang, S. Automated process monitoring in 3D printing using supervised machine learning. *Procedia Manuf.* **26**, 865–870 (2018).
- Jin, Z., Zhang, Z., Demir, K. & Gu, G. X. Machine learning for advanced additive manufacturing. *Matter* **3**, 1541–1556 (2020).
- Jain, A., Armstrong, C. D., Joseph, V. R., Ramprasad, R. & Qi, H. J. Machine-guided discovery of acrylate photopolymer compositions. *ACS Appl. Mater. Interfaces* **16**, 17992–18000 (2024).
- Shukla, S. S., Kuenneth, C. & Ramprasad, R. Polymer informatics beyond homopolymers. *MRS Bulletin* **49**, 17–24 (2024).
- Phan, B. K. et al. Gas permeability, diffusivity, and solubility in polymers: simulation-experiment data fusion and multi-task machine learning. *npj Comput. Mater.* **10**, 186 (2024).
- Ye, X. & Sridhar, T. Effects of the polydispersity on rheological properties of entangled polystyrene solutions. *Macromolecules* **38**, 3442–3449 (2005).
- Boudaghi-Khajehnoob, M., Edwards, B. J. & Khomami, B. Effects of chain length and polydispersity on shear banding in simple shear flow of polymeric melts. *Soft Matter* **16**, 6468–6483 (2020).
- Nichetti, D. & Manas-Zloczower, I. Viscosity model for polydisperse polymer melts. *J. Rheol.* **42**, 951–969 (1998).
- Li, H. et al. Analyzing and predicting the viscosity of polymer nanocomposites in the conditions of temperature, shear rate, and nanoparticle loading with molecular dynamics simulations and machine learning. *J. Phys. Chem. B* **127**, 3596–3605 (2023).
- Karki, B., Bhattarai, D. & Zhang, J. Modeling viscosity of silicate melts using machine learning techniques. AGU Fall Meeting Abstracts MR45A-0078 (2021).
- Takada, S., Suzuki, T., Takebayashi, Y., Ono, T. & Yoda, S. Machine learning assisted optimization of blending process of polyphenylene

- sulfide with elastomer using high speed twin screw extruder. *Sci. Rep.* **11**, 1–9 (2021).
28. Abdullah, M. B. et al. Physics-based and data-driven polymer rheology model. *SPE J.* **28**, 1857–1879 (2023).
 29. Pun, G., Batra, R., Ramprasad, R. & Mishin, Y. Physically informed artificial neural networks for atomistic modeling of materials. *Nat. Commun.* **10**, 1–10 (2019).
 30. Zhang, Z. & Gu, G. X. Physics-informed deep learning for digital materials. *Theor. Appl. Mech. Lett.* **11**, 100220 (2021).
 31. Bradford, G. et al. Chemistry-informed machine learning for polymer electrolyte discovery. *ACS Cent. Sci.* **9**, 206–216 (2023).
 32. Mattheakis, M., Protopapas, P., Sondak, D., Di Giovanni, M. & Kaxiras, E. Physical symmetries embedded in neural networks. Preprint at <https://arxiv.org/abs/1904.08991> (2019).
 33. Raissi, M., Perdikaris, P. & Karniadakis, G. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).
 34. Otsuka, S., Kuwajima, I., Hosoya, J., Xu, Y., Yamazaki, M. PoLyInfo: Polymer Database for polymeric materials design. In: *2011 International Conference on Emerging Intelligent Data and Web Technologies 22–29* (2011).
 35. Rohatgi, A. Webplotdigitizer: Version 4.6 (2022). <https://automeris.io/WebPlotDigitizer>.
 36. Wypych, G. *Handbook of Polymers* 2nd edn (ChemTec Publishing, 2016).
 37. Lewis, P. R. in *Forensic Polymer Engineering* 2nd edn (ed. Lewis, P. R.) 33–69 (Woodhead Publishing, 2016).
 38. Polychronopoulos, N. D., Vlachopoulos, J. in *Functional Polymers* (eds. Jafar Mazumder, M. A., Sheardown, H. & Al-Ahmed, A.) 133–180 (Springer International Publishing, 2019).
 39. Mours, M. & Winter, H. in *Experimental Methods in Polymer Science* (ed. Tanaka, T.) 495–546 (Academic Press, 2000).
 40. Rahaman, N. et al. On the spectral bias of neural networks. In *International Conference on Machine Learning* (pp. 5301–5310), (PMLR, 2019).
 41. Hieber, C. & Chiang, H. Shear-rate-dependence modeling of polymer melt viscosity. *Polym. Eng. Sci.* **32**, 931–938 (1992).
 42. Vlachopoulos, J., Polychronopoulos, N. & Kontopoulou, M. in *Applied Polymer Rheology: Polymeric Fluids with Industrial Applications* (ed. Kontopoulou, M.) 1–27 (Wiley, 2011).
 43. Rubinstein, M. & Colby, R. H. *Polymer Physics* (Oxford University Press, 2003).
 44. Shaw, M. T. On finding the zero-shear-rate viscosity of polymer melts. *Polym. Eng. Sci.* **61**, 1166–1178 (2021).
 45. Hieber, C. A. & Chiang, H. H. Shear-rate-dependence modeling of polymer melt viscosity. *Polym. Eng. Sci.* **32**, 931–938 (1992).
 46. Takahashi, H., Matsuoka, T. & Kurauchi, T. Rheology of polymer melts in high shear rate. *J. Appl. Polym. Sci.* **30**, 4669–4684 (1985).
 47. Helleloid, G. On the computation of viscosity-shear rate temperature master curves for polymeric liquids. (2001).
 48. Ashter, S. A. in *Thermoforming of Single and Multilayer Laminates* (ed. Ashter, S. A.) 123–145 (William Andrew Publishing, 2014).
 49. Paszke, A. et al. In *Advances in Neural Information Processing Systems* 32 (eds Wallach, H. et al.) 8024–8035 (Curran Associates, Inc., 2019).
 50. Pedregosa, F. et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
 51. Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A. & Talwalkar, A. Hyperband: A novel bandit-based approach to hyperparameter optimization. *J. Mach. Learn. Res.* **18**, 1–52 (2018).
 52. Liaw, R. et al. Tune: a research platform for distributed model selection and training. Preprint at 1807.05118 (2018).

Acknowledgements

This work was financially supported by the Office of Naval Research (ONR) through Grant N00014-21-1-2258 and the National Science Foundation (NSF) DMREF Grant 2323695. A.J. acknowledges S. Shukla for help with the fingerprinting of polymer blends. A.J. acknowledges L. Chen and C. Kueneth for valuable discussions on benchmarking ML models.

Author contributions

A.J. is the main architect of the models and dataset and wrote the paper. R.G. helped support the writing and technical soundness of this work. A.R. supported the data collection and analysis of the initial models. J.Q. supported the soundness of the work as applied to polymer manufacturing. R.R. conceived the project and guided the work.

Competing interests

R.R. is a founder of Matmerize, Inc., a company specializing in materials informatics software and services. The other authors have no conflicts of interest to declare.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-025-01532-6>.

Correspondence and requests for materials should be addressed to Rampi Ramprasad.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025