Application of Supervised Machine Learning Regression Algorithms to Prediction of Dielectric Properties of PPy/Kufeki Stone Composites for Energy Implementations

Onder EYECIOGLU[1] 🆔 , Yasar KARABUL[2] 🆔 , Mehmet KILIC[2] 🆔 , Zeynep GUVEN OZDEMIR[2,*] 🆔

[1]*Department of Civil Engineering, Faculty of Engineering and Architecture, Nisantasi University, Istanbul, Turkey*
[2]*Department of Physics, Faculty of Arts and Science, Yildiz Technical University, Istanbul, Turkey*

Highlights

- Supervised machine learning models predicted the dielectric parameters of polypyrrole composites.
- Limited experimental data of the samples were used as training sets for machine learning models.
- The accuracy of the estimated parameters was evaluated by R^2, RMSE, and MAE errors.
- The potential of these composites in energy storage applications was revealed by machine learning.

Article Info	Abstract
	The present study deals with the application of the supervised machine learning regression algorithms known as Linear Regression (LR), Support Vector Machine (SVM), and Gaussian process regression (GPR) to the frequency and temperature-dependent dielectric parameters of polymer/inorganic film composites. The frequency and temperature-dependent experimental data set of the dielectric parameters (ε' and ε'') of Polypyrrole/Kufeki Stone (PPy/KS) has been utilized. ML models were compared based on their model performance and the most suitable was chosen. After choosing the most suitable ML model, at first, the predictions of the same dielectric parameters of the same samples for different temperatures have been made. Then, the predictions of temperature and frequency-dependent ε' and ε'' have been performed for the new PPy based composites consisting of different KS additives that were not produced experimentally. As a result of machine learning, the saturation for KS reinforcing material weight % for dielectric parameters has been determined for capacitor applications. In the light of experimental data and the estimations made by the GPR algorithm, some specific KS additive percentage, working temperature, and frequency ranges have been suggested for the capacitor applications of PPy.

1. INTRODUCTION

In recent years, the production of new composite materials and the determination of their dielectric properties have become very important for energy storage applications. In this respect, the collaboration of machine learning experts and material scientists can provide a great advantage in terms of both time and cost. Then, by producing a few numbers of composite samples and measuring their dielectric properties experimentally, the dielectric performances of the similar series of composites can be predicted for energy storage applications with the help of machine learning (ML) algorithms. As is known, ML is an application of artificial intelligence in which computers estimate the outcomes automatically. The Linear Regression (LR), Support Vector Machine (SVM), and Gaussian process (GP) are the most utilized approaches to represent nonlinear relationships between pairs of random variables for both regression and classification types. In this manner, LR, SVM and GP regression (GPR) algorithms have been tried to predict the dielectric parameters of the polymer/inorganic film composites with high accuracy. While LR is a regression machine learning algorithm that focuses on finding the relationship between two continuous variables, SVMs are a set of supervised learning methods used for classification, regression, and outliers detection. Additionally, GPR is evolved from a kernel function that determines the similarity between structures [1].

---

*Corresponding author, e-mail: zguven@yildiz.edu.tr

As is known, supervised or unsupervised ML algorithms have been applied to predict many properties in different disciplines including physics, materials science, signal processing, weather forecast, medical science, cybersecurity, mechanical engineering, etc. Among these applications, the integration of ML algorithms into material science has enabled the electrical parameters of composite series to be estimated with the help of a small number of experimental data, thus accelerating the discovery of new materials. In this case, by using the experimental values of the electrical parameters of composite series as a training data set for ML, the electrical data of similar and unmanufactured composite series can be estimated. In this way, it is possible to propose the electrical parameter limits for the composite materials which have not been produced yet and to determine their possible technological application areas by saving time and production costs to a great extent. Although precisely material production and characterization are indispensable for the discovery of new composite materials, the estimation of specific composite component percentages, working temperature and frequency range for desirable and adjustable electrical properties will make a significant contribution to material science. Considering ML algorithms, it is possible to estimate a given material's features including its structure [2,3], molecular properties [4,5], bandgaps [6,7], and dielectric properties [8]. Dielectric properties of a material can also be estimated by the artificial neural network method [9-11].

When the ML studies focused on material science and condensed matter are evaluated, some studies are remarkable. For example, Pilania et al. attempted to find an accurate and effective way to profile a given material by considering an ML method based on the fingerprint vectors depending on configurational and compositional knowledge or the electronic charge density distribution of the material [8]. Pilania et al. also researched the estimations of the electronic bandgap of perovskite materials by using ML. Mannodi-Kanakkithodi et al. have also used ML method for designing polymer-based dielectrics. They tried to obtain a fast decision for a given polymeric material either it is suitable for dielectric applications or not by using statistical learning models [12]. Moreover, Schütt et al. proposed an ML-based algorithm for the quick estimation of the states' density of materials at the Fermi energy level by using a training data set consist of local spin-density computations [13]. If it is focused on the recent studies examining different regression models in terms of condensed matter and materials science, it is seen that the studies are generally clustered in LR, SVM, and GPR models. For example, Schmidt et al. have evolved a new method by using linear models to obtain ground-state electron densities of binary alloys with R^2 values higher than 0.999 [14]. On the other hand, Owolabi et al. used the SVM algorithm to predict the superconducting transition temperature (T_c) of the Fe-based superconductors at ambient pressure. They predicted not only the T_c with high accuracy of over 99% but also they could suggest a strong relationship between the T_c and the lattice parameters by SVMs [15]. Additionally, Pote and Melko studied the availability of the usage of SVMs on the prediction of order parameters and Hamiltonian constraints for 2D-spin systems. In this context, they emphasized the interpretable classification ability of SVMs as well as their high performance on the property detection of the data sets of condensed matter and many-body systems [16]. Another condensed matter and also nanotechnology application of ML has been carried out by Caro-Gutiérrez et al. in 2020. They used SVMs to predict the average length of vertically aligned titanium dioxide nanotubes that are utilized in different applications including electronics, biomedical implements, and environmental industries. They have reached the lowest error in prediction for the linear SVM for regression [17]. On the other hand, Zhang and Xu have estimated the Curie temperature of magnetocaloric lanthanum manganites by using GPR with a high accuracy degree. They studied approximately one hundred lattices of the lanthanum manganites and they have succeeded to develop a statistical correlation between the Curie temperature and lattice parameters [18]. In addition to these studies, Raimbault et al. applied the GPR algorithm for the estimation of the molecular crystals' vibrational Raman spectra. They showed that the related materials' dielectric susceptibility and static polarizability can be predicted accurately by GPR [1]. By referring to the extensive usage of LR, SVM, and GPR in condensed matter physics, the related ML algorithms have been preferred to try to predict the energy storage performance of the Polypyrrole (PPy) based Kufeki Stone (KS) added composites.

As is known, PPy is a conductive polymer that has both a difficult and expensive manufacturing process. To determine the usage of PPy and also its composites in energy storage applications, the frequency- and temperature-dependent measurements have to be performed. On the other hand, both the synthesis of the composites and these experiments have high costs and require a long time. Due to overcome these

advantages, the suitable supervised ML algorithm for dielectric parameters has been chosen among linear regression, support vector machine, and Gaussian process regression and then the limits of such a valuable material such as the appropriate frequency range, operating temperature, and natural material additive percentage for electronic applications have been determined in the present manuscript.

2. THEORETICAL

2.1. Complex Permittivity of Dielectric Materials

The charge storage performance of dielectric material is characterized by the real and imaginary parts of complex permittivity (ε^*). The real part (ε') is responsible to represent the charge storage ability of a dielectric material.

$$\varepsilon' = \frac{Cd}{\varepsilon_0 A} \qquad (1)$$

where C is the capacitance of the capacitor with the electrode surface area of A and separation length of d when it is filled with a dielectric material. The imaginary part of complex permittivity (ε'') is also an important parameter to determine a dielectric material's performance and it is related to the dielectric loss. As is known, the imaginary part of complex permittivity represents an energy loss when the polarization lags the external electric field and it may occur due to interfacial polarization, dc conductivity, or movement of dipoles [19].

Additionally, it is possible to construct a relation between the real and imaginary parts of ε^*:

$$tan\delta = \frac{\varepsilon''}{\varepsilon'} \qquad (2)$$

where $tan\delta$ is the tangent loss.

2.2. Machine Learning Models

Machine learning (ML) is a computational methodology of computer science that can be used to interpret an experimental dataset, to drive meaning from this dataset, and then to optimize the performance criteria used in any problem-solving procedures. The ML algorithms are divided into two main categories, unsupervised and supervised algorithms that can be used for both classification (for discrete outputs) and regression (for continuous output). In this study, Linear Regression (LR), Support Vector Machine (SVM) regressor, and Gaussian process regression (GPR) algorithms have been used to predict the frequency and temperature-dependent dielectric parameters of polymer/inorganic film composites.

2.2.1. Linear regression (LR) model

Linear regression is one of the simplest models of both machine learning and statistical. It is a well-known machine learning algorithm based on data-driven supervised learning. In this model, it is assumed that a linear function ($y(\{x_n\})$) identifies the relationship between observations $\{x_n\}$ and target values $\{t_n\}$ which are components of the training data sets [20]. To express the uncertainty about the value of t for each value of x, a predictive distribution $p(t|x)$ should be modeled.

The simplest linear model for linear regression is one that involves a linear combination of the input variables,

$$y(x) = w_0 + \sum_{i=1}^{n-1} w_i \phi_i(x) \qquad (3)$$

here, w_i, and ϕ_i are weight parameters and basis functions correspondingly. The w_0 parameter is the bias parameter that allows any fixed offset in the data. It is possible to define many kinds of basic function such as $\phi_i(x) = x_i$. We need a loss function to measure the error the predicted values and the actual values

$$\mathcal{L}(x) = \sum_{i=1}^{n}(y_i - p_i)^2 .\tag{4}$$

In the least-squares approach, the main challenge is to determine the parameters w_0 and w_i, that give the minimum error for target values

$$\min_w \|y_i - p_i\|^2 .\tag{5}$$

2.2.2. Support vector machine (SVM) model

Another well-known non-linear machine learning tool for both classification and regression is the Support vector machine (SVM). SVMs are versatile models that can perform linear or nonlinear classification and regression. The Support Vector Regression (SVR) uses the same principles as the SVM for classification. SVM needs only a subset of this training instances, called support vector (SV) for predicting the future, once the training data set has been trained and determined using the model parameters.

The main aim of the SVM is to construct a separating hyperplane that correctly separates two classes with a maximum margin in high dimensional feature space by linear or non-linear mapping functions. SVM generalization to SVR is accomplished by introducing an ε-insensitive loss function.
Suppose training vectors $\mathcal{D} = \{(x_i, y_i), i = 1,2,..,N\}$, $x \in \mathbb{R}^N, y \in \mathbb{R}$ where (x_i is a multivariate set of N observations with observed response values y_i. For simplicity, consider a linear function describe separating hyperplane

$$y = f(x) = \langle w, x \rangle + b = \sum_{i=1}^{N} w_i x_i + b, \quad x, w \in \mathbb{R}^N, y, b \in \mathbb{R} .\tag{6}$$

SVR attempts to find the narrowest tube centered around the surface while minimizing the prediction error

$$\min_w \frac{1}{2} \|w\|^2 .\tag{7}$$

The goal is to minimize the error between the predicted value of the function and the actual output for a given input. SVR accepts the *ε-insensitive loss function* and eliminates predictions further from e from the desired output

$$L_e = \big(y, f(x,w)\big) = \begin{cases} 0 & |y - f(x,w)| \leq \varepsilon \\ |y - f(x,w)| - \varepsilon & otherwise \end{cases} .\tag{8}$$

The optimal regression function is given by the minimum of the functional;

$$\Phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i^- + \xi_i^+ \tag{9}$$

where C is the box constraint and ξ_i^- and ξ_i^+ are slack variables representing upper and lower constraints on the outputs of the system, respectively. It can be simpler to solve the optimization problem in its Lagrange dual formulation

$$L(\alpha) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)x_i'x_j + \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{N} y_i(\alpha_i - \alpha_i^*).\tag{10}$$

α_i and α_i^* are The Lagrange multipliers, or dual variables which are nonnegative real numbers. Based on the Karush-Kuhn-Tucker (KKT) conditions, the minimum of the optimization problem can be found by taking partial derivative with respect to dual variables. So, the parameter w can be described as;

$$w = \sum_{i=1}^{N}(\alpha_i^* - \alpha_i)x_i. \tag{11}$$

For the case that the problems cannot be described using linear models, the Lagrange dual formulation is extended to non-linear functions. For non-linear functions, data can be mapped to a higher dimensional space called kernel field to achieve higher accuracy. A nonlinear SVR model is obtained by replacing the dot product $x_i' x_j$ with a nonlinear kernel function $k(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle$, where $\varphi(x_i)$ is the transformation from feature to kernel space. By minimizing the Lagrange dual formulation for non-linear function

$$L(\alpha) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)k(x_i, x_j) + \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{N}y_i(\alpha_i - \alpha_i^*) \tag{12}$$

$$w = \sum_{i=1}^{N}(\alpha_i^* - \alpha_i)\varphi(x_i). \tag{13}$$

The function used to predict new values is equal to

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)k(x_i, x) + b . \tag{14}$$

In this study, the Gaussian kernel function has been used

$$k(x_i, x) = e^{(-\|x_i - x\|^2)} . \tag{15}$$

2.2.3. Gaussian process regression (GPR) method

GPR which is a kernel-based, non-parametric supervised ML method, can be used to apply the Bayesian approach for regression problems [21].

Let $D = \{(x_i, y_i) | i = 1,2, \dots, n\}$, is the input dataset of n observable. The conditional distribution $p(x|y)$ depicts the dependency of an observable or dependent variable $y \in R$ on the corresponding set of independent variables $x \in R^m$. This dependency which can be presented by a latent function $f: X \to R$, is of form

$$p(y|f(x), \theta) = N(y|f(x), \sigma^2) \tag{16}$$

where N is noise function and σ^2 is error variance.

For the non-parametric model approach, the latent function $f: X \to R$ should be expressed in terms of a prior distribution. "*A Gaussian Process (GP) is a collection of random variables, any finite number of which have (consistent) joint Gaussian distributions*" [22]. The GP is a generalization of the Gaussian distribution and defined by its mean function $(m(x))$ and a covariance function $(k(x, x'))$

$$f \sim G(m(x), k(x, x')). \tag{17}$$

The mean of the latent dependency function defined as, $m(x) = E[f(x)]$.

The key factor of GP is the covariance function (or kernel) that controls the properties of the GP. The covariance function is defined as

$$k(x, x') = cov[f(x), f(x')] = E[(f(x) - m(x))(f(x') - m(x'))]. \tag{18}$$

While the mean function can be selected arbitrarily, it is a crucial problem that is to find suitable properties for the covariance function.

In this study, the Matern 5/2 kernel function has been utilized used as a covariance function. The Matern 5/2 covariance function is defined as;

$$k(x, x') = \sigma_f^2 \left(1 + \frac{\sqrt{5}r}{\sigma_l} + \frac{5r^2}{3\sigma_l^2}\right) exp\left(-\frac{\sqrt{5}r}{\sigma_l}\right) \tag{19}$$

where $r = \sqrt{(x - x')^T(x - x')}$ (Euclidian distance) and σ_l is the characteristic length scale, and σ_f is the signal standard deviation [21, 23].

2.3. Error Measurements

In this study, coefficient of determination (R^2), root mean square (RMS) and mean absolute error (MAE) metrics have been used to measure the performance of the mentioned machine learning regression models. These metrics are defined in Equations (20), (21) and (22):

$$R^2 = 1 - \frac{\sum_{j=1}^{N}(y_j - \hat{y}_j)^2}{\sum_{j=1}^{N}(y_j - \bar{y})^2} \tag{20}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(y_j - \hat{y}_j\right)^2} \tag{21}$$

$$MAE = \frac{\sum_{j=1}^{N}|y_j - \hat{y}_j|}{n} = \frac{\sum_{j=1}^{N}|e_i|}{n} \tag{22}$$

where y_j and $\hat{y}_j$ are the actual response and predicted response of observation values respectively, and $\bar{y}$ is the average value of the validation dataset.

3. RESULTS

3.1. Experimental Results for Machine Learning

PPy/KS composites have been prepared by the following steps given in Figure 1(a). As is illustrated in Figure 1(a), to obtain PPy/KS composites PPy and KS powders have been mixed in appropriate ratios and ground for 20 min. in an agate mortar. Then the mixtures have been pressed into pellets by applying 4000 psi pressure at room temperature for 5 min. The pellets with 10 mm diameter had thicknesses varying between 1 and 2 mm. Then the samples have been placed between two gold electrodes in a measurement cell. The experimental dielectric data set of the samples have been obtained by NOVO Control Broadband Dielectric/Impedance analyzer with Quatro Cryosystem between 1Hz–40MHz at different temperatures [24]. By using some of the experimental data of the samples as the training set, the different supervised ML models have been performed by following the steps given in Figure 1(b).
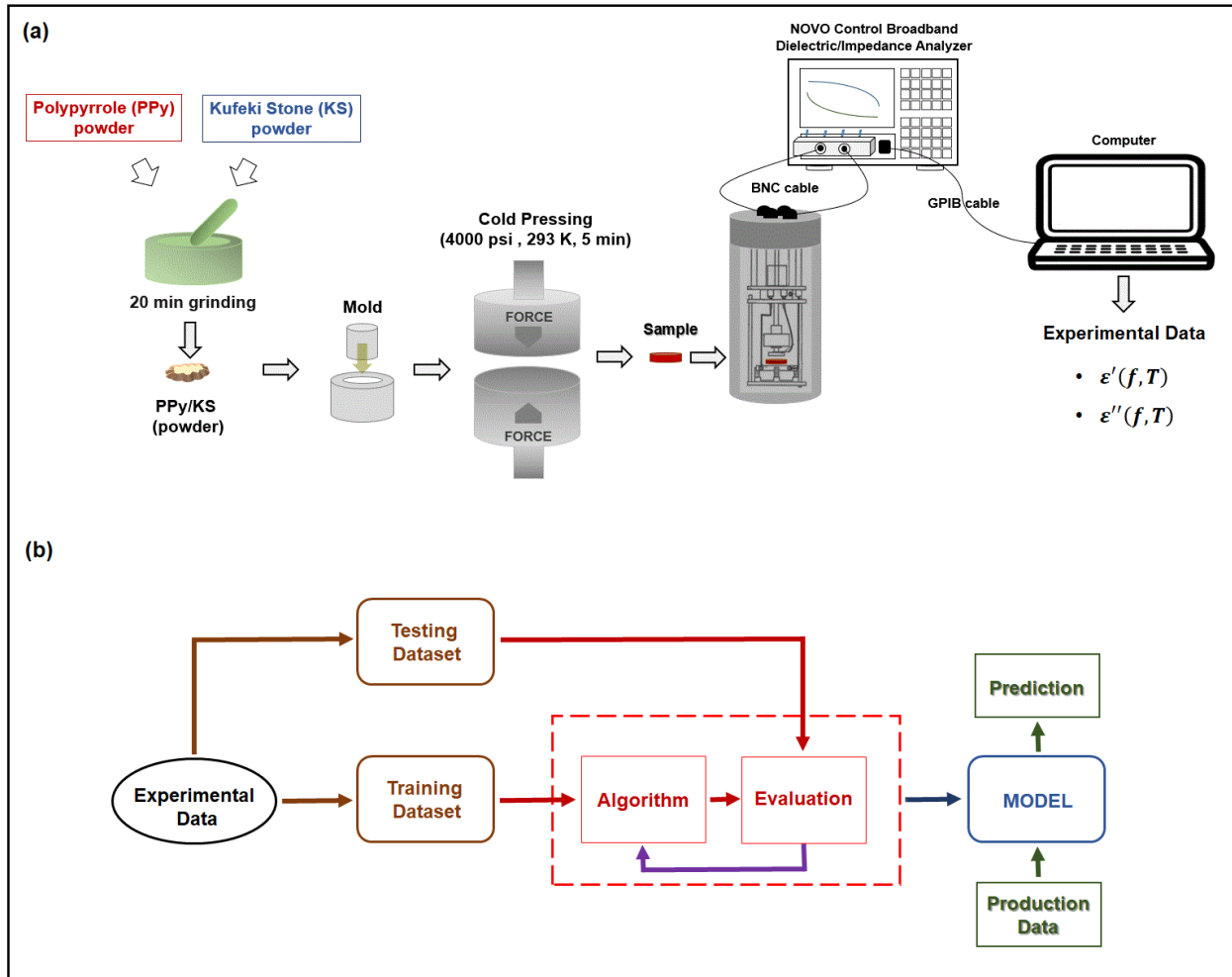
Figure 1. *The schematic diagram of the steps followed to obtain a) experimental and b) ML workflow diagram for the prediction of dielectric data*

The experimental results have been obtained for pure PPy along with 5% KS, 10% KS, and 20% KS added PPy composites for 296, 313, 333, and 353 K. In each experimental data, the parameters has been measured for 68 different frequency values at each temperature. As an example, the variations of the ε' and ε'' with the frequency of two of the samples (pure PPy and PPy/10% KS), which were used as experimental data in machine learning, have been shown in Figure 2.
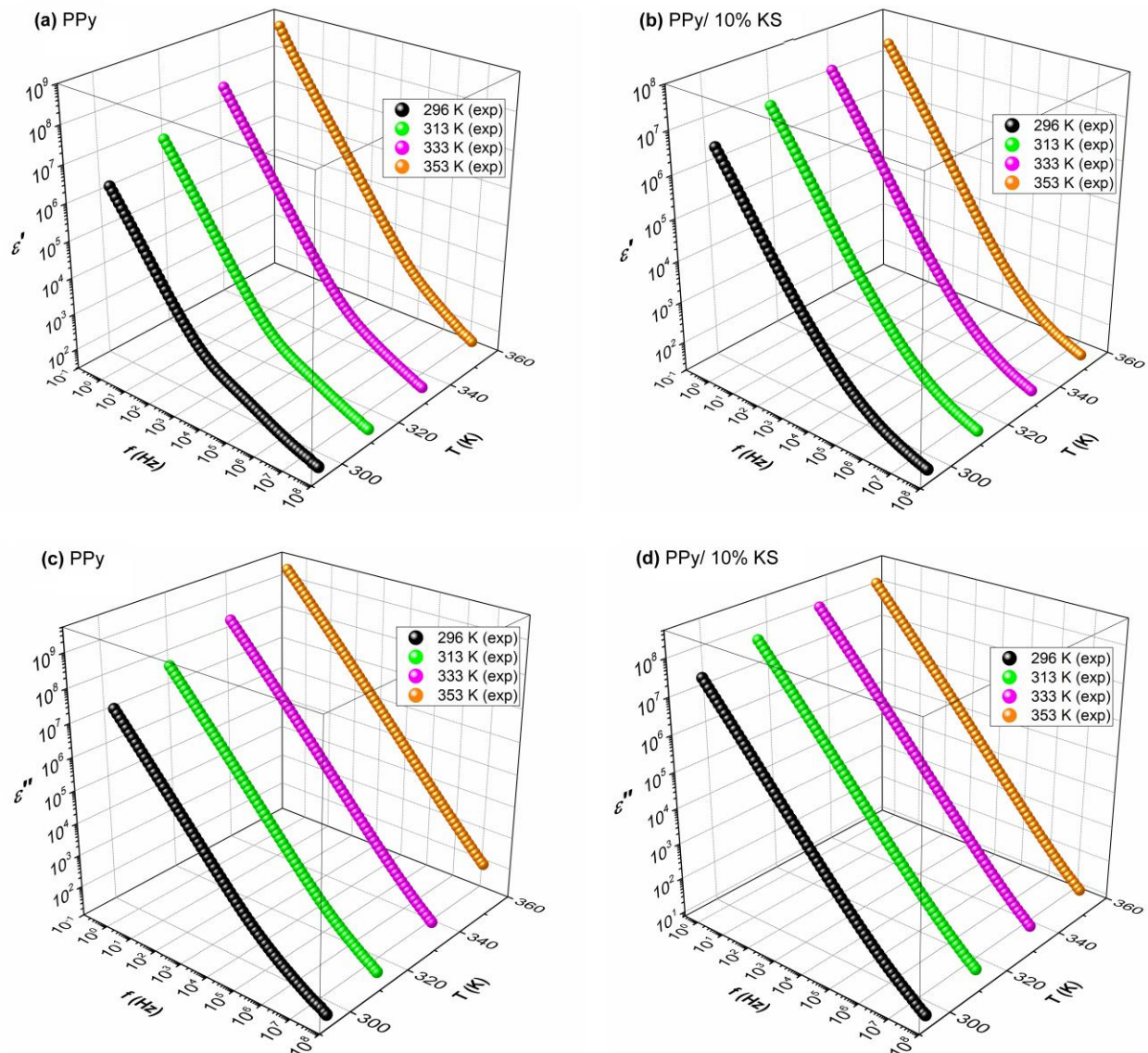
Figure 2. *The frequency dependence of the real part of the complex permittivity for a) PPy and b) PPy/ 10% KS for 296, 313, 333, and 353 K. The frequency dependence of the imaginary part of the complex permittivity for c) PPy and d) PPy/ 10% KS for 296, 313, 333 and 353 K*

As shown in Figures 2 (a) and (b), the real parts of the complex permittivity of the samples show strong frequency dependence and exhibit a significant decrease with increasing frequency. In this context, the addition of KS to PPy did not change the frequency dependence of PPy. Additionally, it has been determined that while the real part of complex permittivity of PPy increases with 10% KS additive for 296 K and 313 K, it decreases with 10% KS addition for 333 K and 373 K. This result has been obtained by comparing the ε' values of PPy and PPy/10% KS composites at some low, mid, and high spot frequencies (1 Hz, 1 kHz, and 1 MHz). On the other hand, both pure PPy and the PPy/10% KS composite increased their real permittivity value hierarchically as the temperature increases from 296 K to 353 K. This behavior can be explained by the increased polarization in the PPy matrix due to the increased number of oriented dipoles. A similar characteristic determined for ε' with temperature and frequency is valid for other samples (See Figure 3 in Section 3.2).

Like the frequency dependence of ε', the imaginary part of the complex permittivity of the samples has a strong dispersion at all temperatures studied. Moreover, an increase in temperature resulted in an increase in the ε'' at all frequencies for both samples. Another important result has also been observed from Figure 2 (c) and (d) which represent the dielectric loss of the materials. The dielectric loss of the PPy decreases with the 10% KS addition at each temperature and frequency.

In the light of the comparison between the real and imaginary components of the complex permittivity of PPy and PPy/10% KS composite, while the charge storage ability of PPy increased with 10 % KS addition, its energy loss decreases. From this point of view, it can be specified that the KS added PPy matrix may have the potential to use in energy storage implementations. In this respect, the optimum KS additive percentage, working temperature and frequency ranges for capacitor applications have been aimed to determine by using machine learning.

3.2. Performance Comparison of Machine Learning Algorithms

In this study, three different Machine Learning models (Linear Regression (LR), Support Vector Machine (SVM) Regression, Gaussian Process Regression (GPR)) have been set to predict unknown frequency, temperature, and KS addition rate-dependent dielectric parameters of polymer/inorganic film composites as supervised ML models. Obtaining the finite input dataset, which is needed by supervised ML, has been explained in section 3.1. This dataset includes 1088 observations with 3 input variables (frequency, temperature, and KS addition rate) and 1 target variable (ε' or ε'') (See Table 1). The mean and standard deviation values have also been given in Table 1. This data set has been divided into 2 parts to use them in training and test steps during ML training and testing procedures. 80% of the dataset has been used as the training data, while the remaining 20% of the total dataset has been utilized as the testing data. Logarithmic scaling has been applied to the input dataset to prepare it for training.

Table 1. *The statistical properties of the total dataset*

Properties	Features				
	f (Hz)	Concentration	T (K)	ε'	ε''
Count	1088	1088	1088	1088	1088
Mean value	2.699×10^6	0.088	323.750	7021.376	39293.610
Standard deviation	7.472×10^6	0.074	21.380	76.148	140.268
Min. value	1	0	296	17.443	4.730
Max. value	4×10^7	0.200	353	7.81×10^8	3.68×10^9

The scatter matrix has been used to understand the relations between the variables and to estimate the covariance. The scatter plots of all features in the input dataset for ε' and ε'' results have been illustrated in Figures 3 and 4, respectively. Diagonals of these matrixes show the covariance between random variables. The covariance relations of the ε' and ε'' indicates that these parameters have Gaussian distribution.
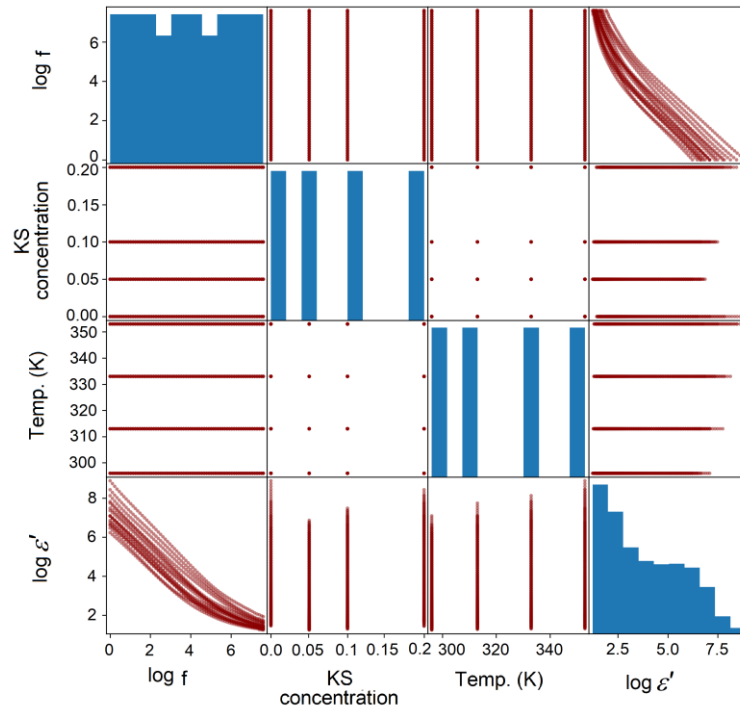
Figure 3. *The scatter plot of all features in the total dataset for* ε'
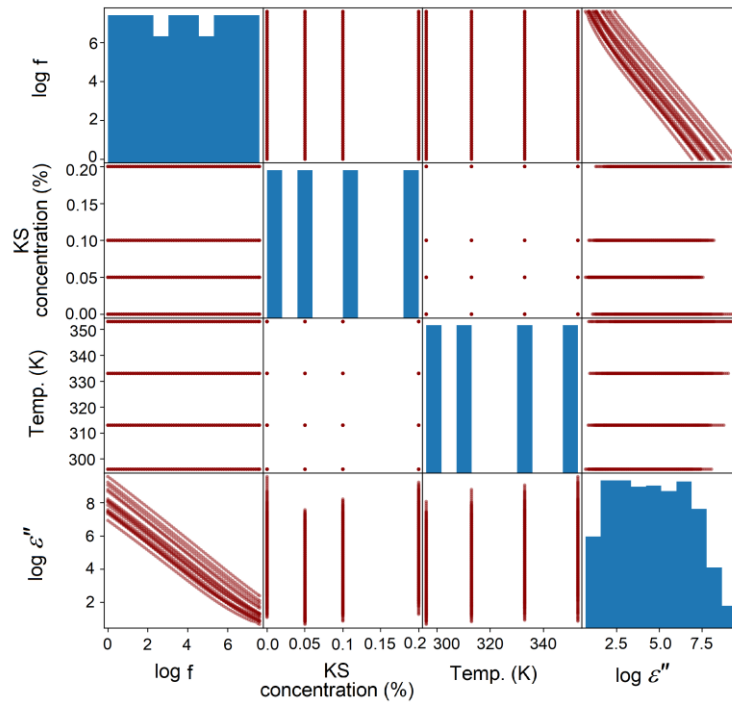
Figure 4. *The scatter plot of all features in the total dataset for* ε''

The predictions of the ε' and ε'' have been realized by using ML models which are kernel-based data-driven supervised ML models. These models have been implemented by employing MATLAB Machine learning toolbox app. The residuals plots shown in Figure 5, have been utilized to analyze the performances of the ML models after the training.
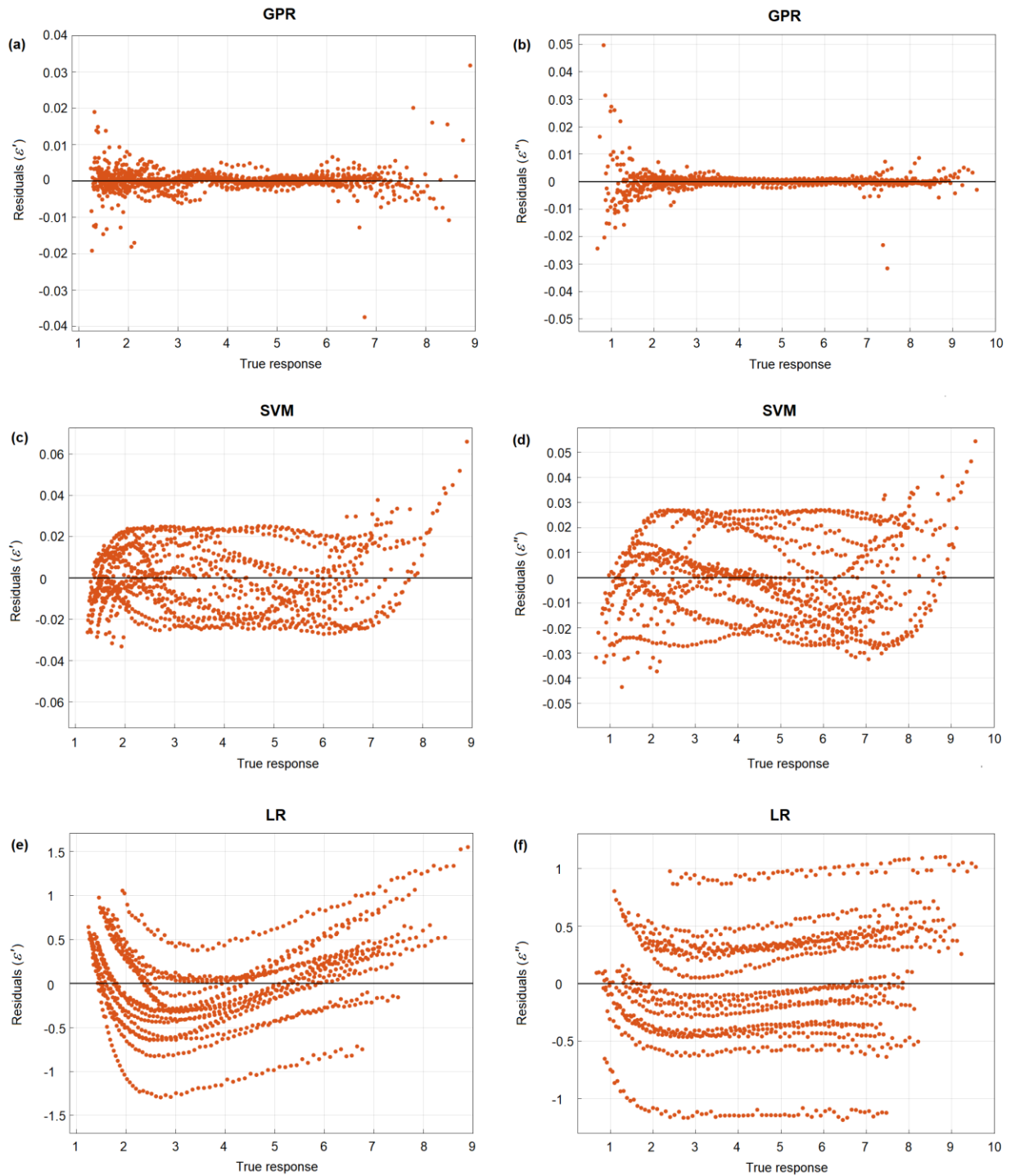
Figure 5. *The residuals plots for the GPR, SVM, and LR models*

The accuracy validations of each ML models have been estimated by a k-fold cross-validation procedure where the value of the k has been chosen as 10. In this procedure, 80% of the input dataset has been used as training data and the remaining 20% data points have been utilized as testing data. The performance of the employed ML models have been compared by using the coefficient of determination (R^2), root mean squared error (RMSE) and mean absolute error (MAE) model statistics metrics. These model statistics metrics for ε' and ε'' have been listed in Table 2.

Table 2. *Error results of the models*

| Model	Model statistics	Parameters	
		ε'	ε''
GPR	R^2	0.99	0.99
	RMSE	0.0380	0.0032
	MAE	0.0019	0.0013
SVM	R^2	0.99	0.99
	RMSE	0.1587	0.1778
	MAE	0.1346	0.1496
LR	R^2	0.94	0.95
	RMSE	0.4714	0.4966
	MAE	0.3725	0.4068

Comparison and model selection has been evaluated by analyzing residuals plots shown in Figure 5, and the performance metrics have been listed in Table 2. Based on Figure 5 and Table 2, it is seen that there are good agreements between Matern 5/2 GPR model predictions results and observed ones. Also, Matern 5/2 GPR model has the best model statistics values among others. According to these results, it has been concluded that GPR is a suitable model to predict ε' and ε'' parameters.

3.3. Estimation Results for Dielectric Parameters of PPy/KS Composites by Gaussian Process Regression

In section 3.2, it has been determined that the estimation of both dielectric parameters of the samples performed by the GPR algorithm gives the best results among other models. From this point of view, the prediction of the dielectric parameters of PPy/KS composites has been realized by GPR in two stages. In the first stage, ε' and ε'' have been predicted for the present samples at experimentally untreated temperatures. In the second stage, the estimation of the related dielectric parameters of the samples which are not prepared experimentally was determined for the temperature interval of 296-353 K and frequency interval of 1 Hz- 40 MHz. In this respect, the GPR results have been given in two subsections.

3.3.1. GPR results for pure PPy and PPy/ 5% KS, PPy/ 10% KS, and PPy/ 20 KS composites at new temperatures

The frequency dependences of the real part of complex permittivity of the prepared samples for the untreated temperatures (303 K, 323 K, and 353 K) have been given in Figure 6. These new temperatures are between experimental temperatures i.e. in the interpolation zone. The experimental and GPR data for ε' have also been shown together in Figure 6. While the experimental data are symbolized with small spheres with different colors depending on the temperature, the GPR estimations are shown by star symbols in Figure 6.
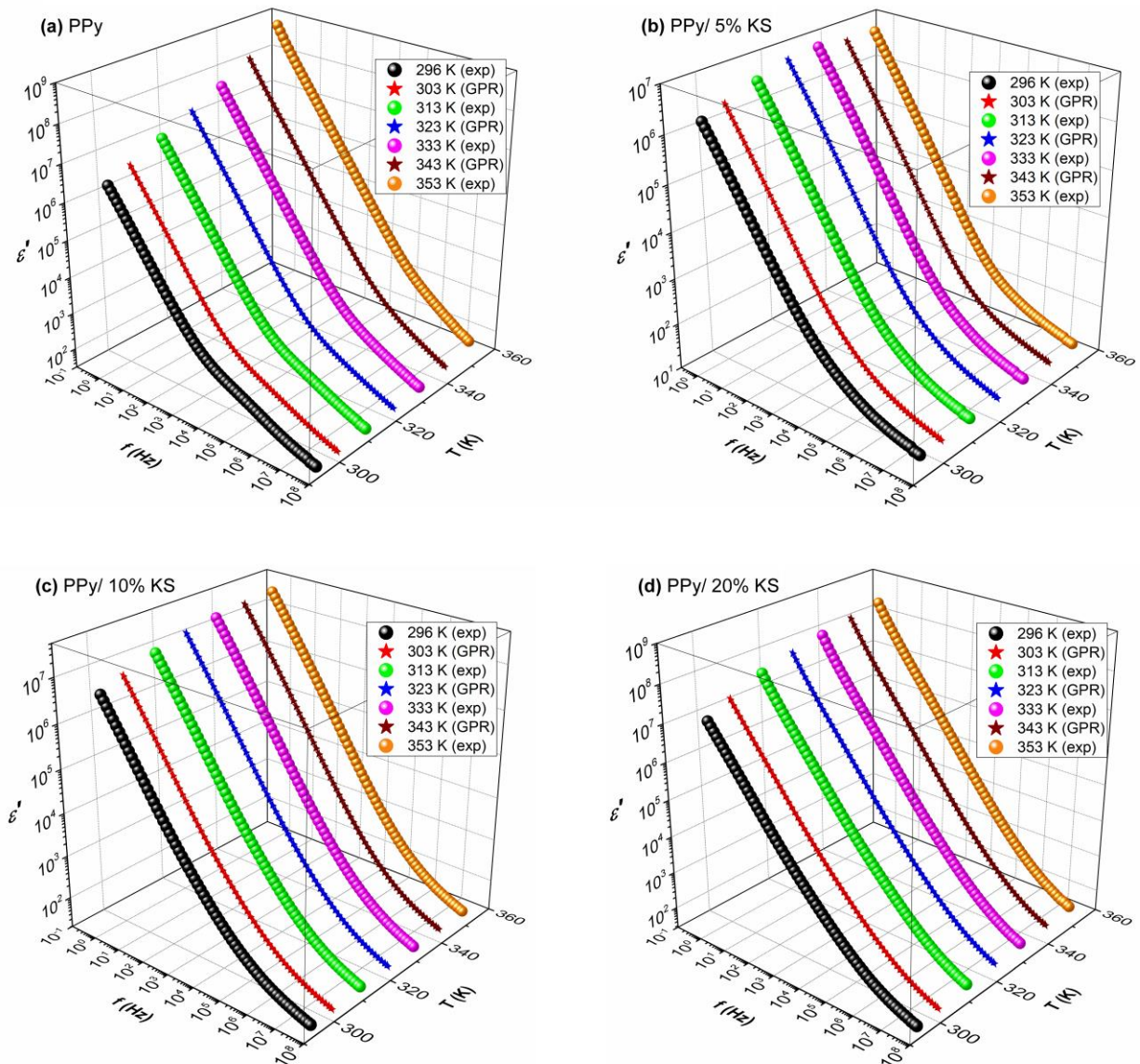
Figure 6. *The variation of the real part of permittivity with frequency for experimentally treated and untreated temperatures of the a) pure PPy, b) PPy/5% KS, c) PPy/10% KS, and d) PPy/20%KS*

As shown in Figure 6, the new ε' data of each sample is increasing with temperature in accordance with the experimental data. Additionally, the general frequency response of the ε' is also conserved at these new temperatures for all samples. From the technological point of view, the GPR method has also been applied to the imaginary part of complex permittivity data of the present samples for these new temperatures. The results have been given with the experimental data to emphasize the compatibility between the experimental and GPR findings in Figure 7.
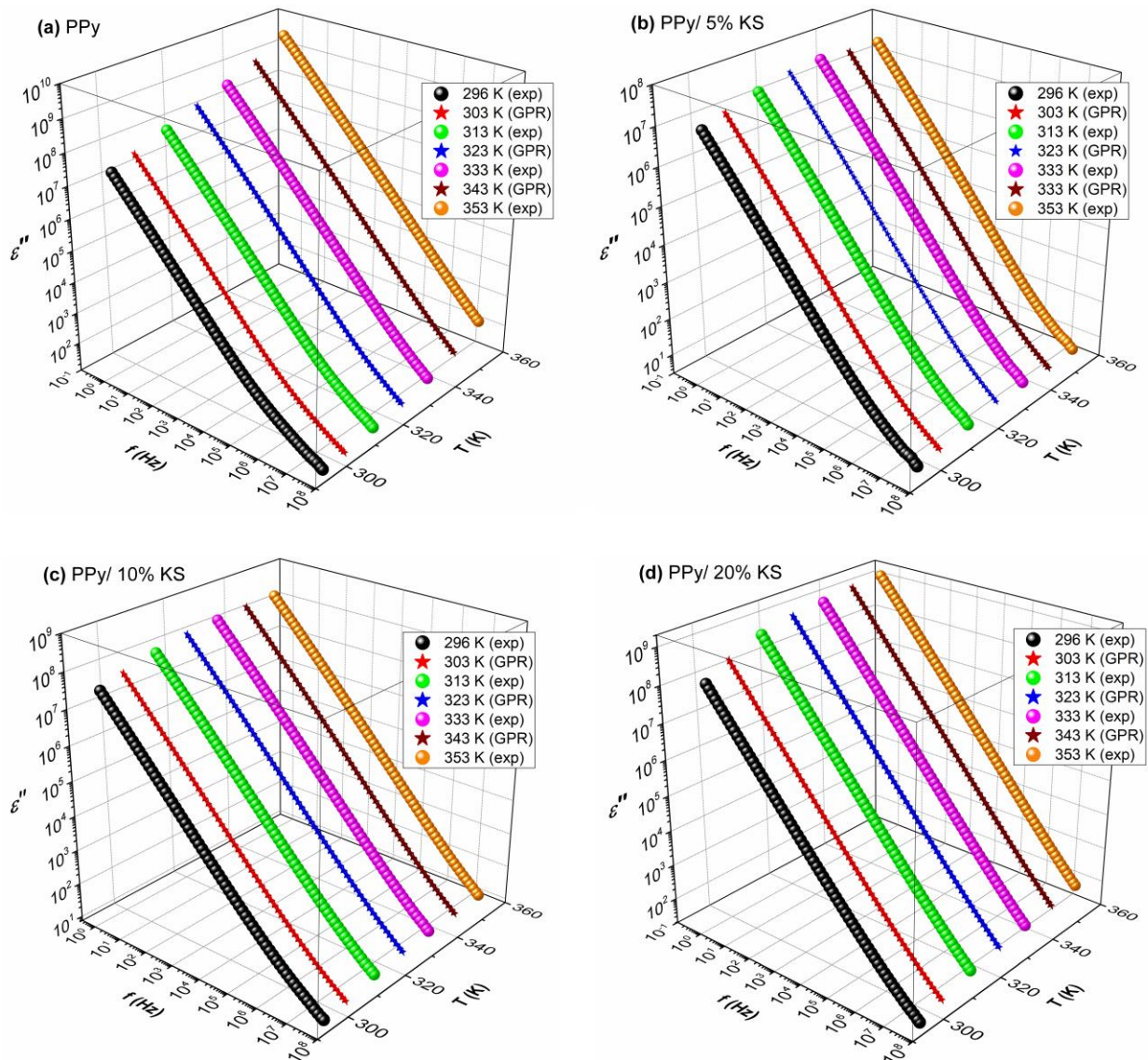
Figure 7. *The variation of the imaginary part of permittivity with frequency for experimentally treated and untreated temperatures of the a) pure PPy, b) PPy/5% KS, c) PPy/10% KS, and d) PPy/20%KS*

When the technological importance of the samples are evaluated from the dielectric loss point of view, it has been observed that while the dielectric loss decreases approximately by 3 times and 120 times at low and high temperatures for the 5% KS addition, it also remains less than that of pure PPy at all temperatures for the 10% KS. But when KS additive reaches 20% in the PPy matrix, the dielectric loss increases approximately by 5 times relative at all temperatures.

If the real and imaginary parts of the complex permittivity values of the present samples are interpreted at the same time, it has emerged that some new KS additive percentages should be tried to determine the PPy based composites which may have high energy storage capability performance. Because the samples with high dielectric constant and low dielectric loss or high dielectric constant and high dielectric loss can be used in capacitor and decoupling capacitor applications, respectively. Therefore, it is technologically important to know which temperature and KS% contributions will result in the highest and lowest ε' and ε'' values. For this purpose, the GPR algorithm whose reliability has been tested in the previous section can be preferred as a zero-cost method for specifying the highest and lowest values of the relevant parameters.

3.3.2. GPR results for unprepared PPy/7.5% KS, PPy/ 15% KS, PPy/ 25% KS and PPy/30 KS composites between 296-353 K

The GPR algorithm has been applied to some new KS% additive in the PPy matrix to determine the limits of ε' and ε'' values. In this section, the frequency dependence of the real and imaginary parts of complex permittivity has been estimated for 7.5%, 15%, 25%, and 30% KS additives at seven different temperatures. The variations of ε' and ε'' of the related experimentally unproduced samples with frequency have been given in Figures 8 and 9, respectively.

When the ε' data shown in Figure 8 are considered together with the data given in Figure 6, it has been observed that the ε' values are still less than pure PPy at all temperatures in the low-frequency region for the 7.5% KS addition. However, as the KS additive concentration increases from 10% to 25% KS, ε' value increases in the low-frequency region for the temperature range of 296-313 K. On the other hand, when the ε' values of 30% KS added PPy composite estimated by GRP are examined, it has been found that the ε' value decreases 10 times at low temperatures and 1000 times at high temperatures compared to pure PPy in the low-frequency region.
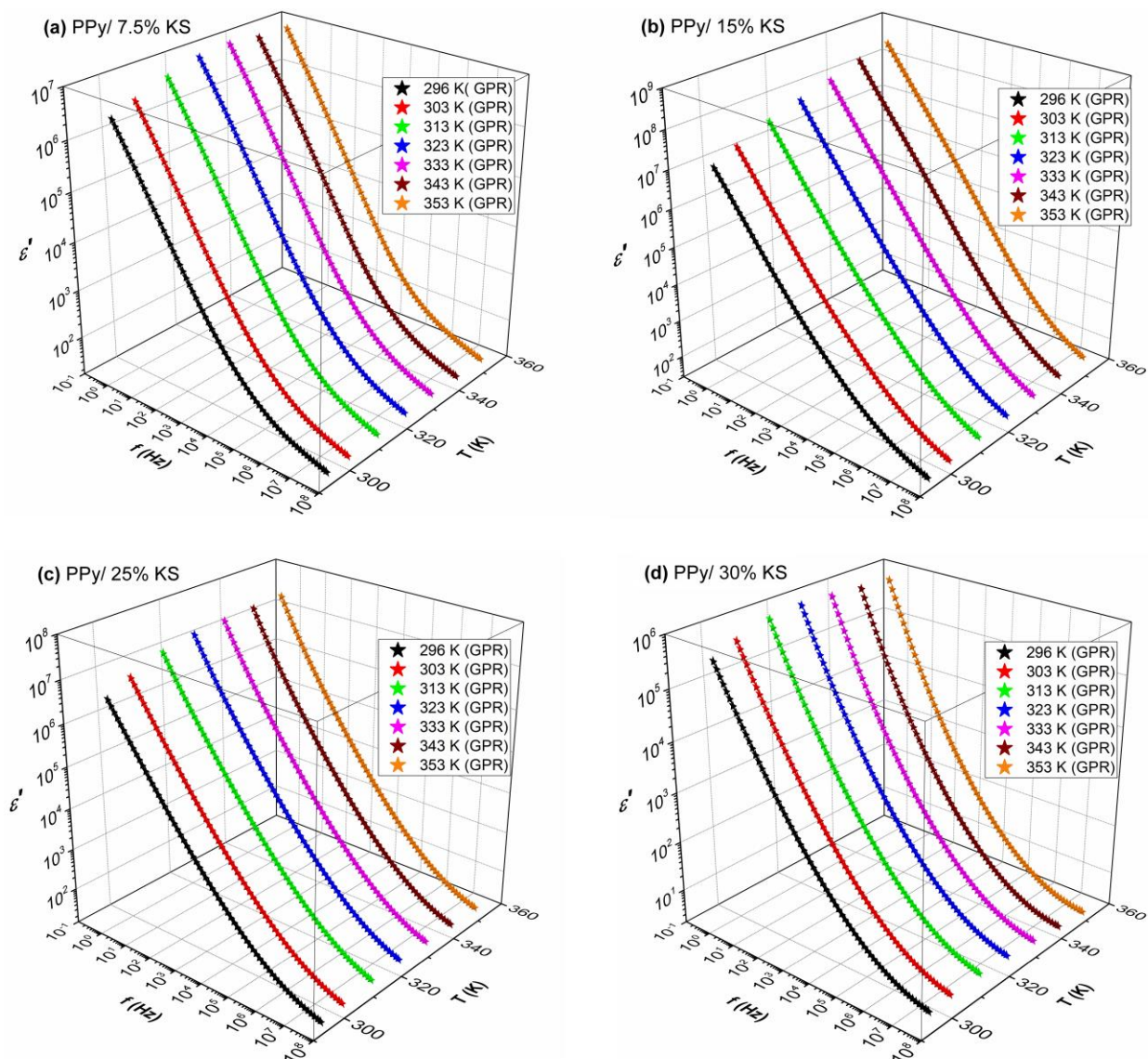
Figure 8. *The variation of the real part of permittivity with frequency for 296-353K of the a) PPy/7.5% KS, b) PPy/15% KS, c) PPy/25% KS, and d) PPy/30%KS*
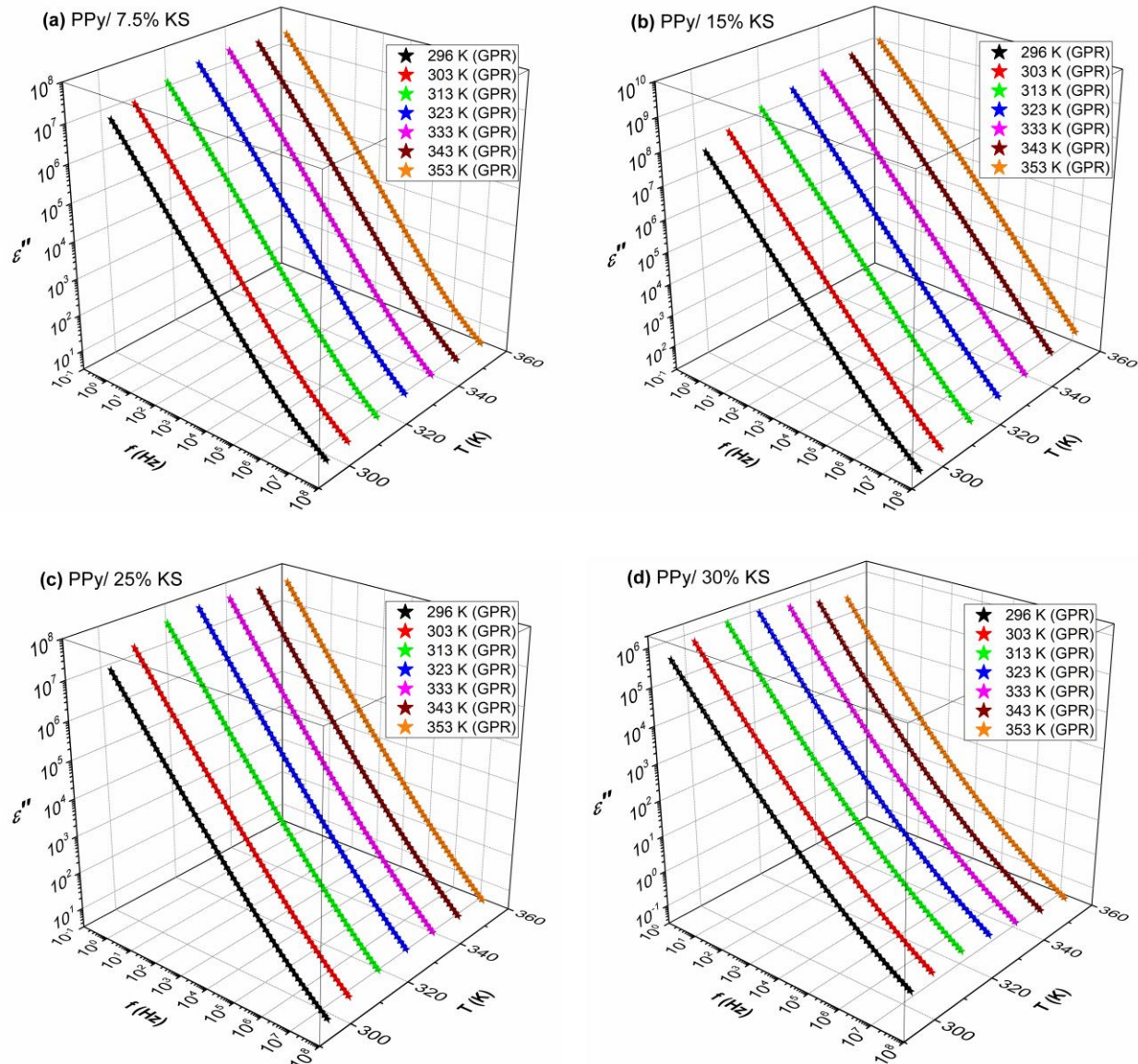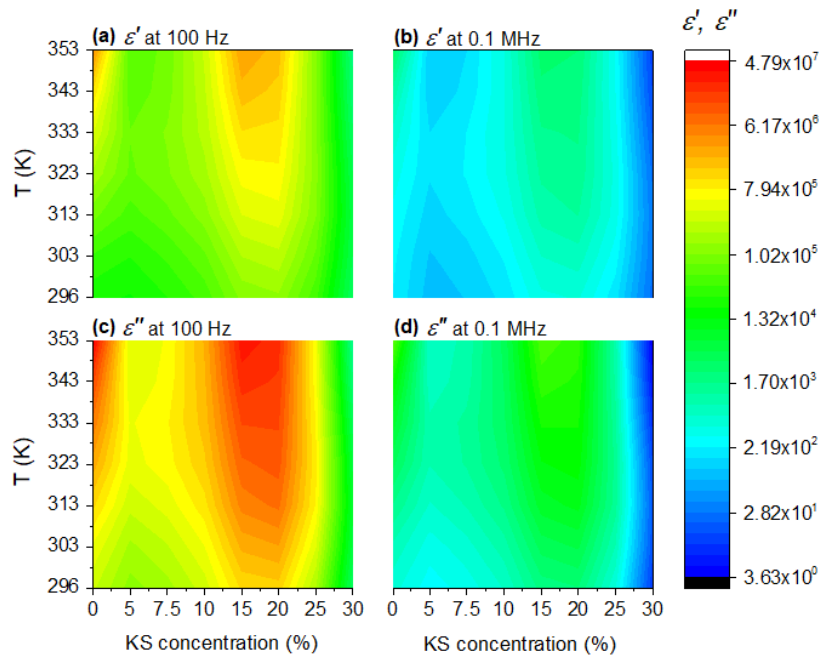
Figure 9. *The variation of the imaginary part of permittivity with frequency for 296-353K of the a)*
PPy/7.5% KS, b) PPy/15% KS, c) PPy/25% KS, and d) PPy/30%KS

If a similar examination is performed for dielectric loss at the low-frequency region depending on the KS contribution percentage, it has been detected that slightly lower ε'' values relative to pure PPy are observed up to 10% KS additive. On the other hand, significant reductions in ε'' have been observed at all temperatures for KS additives above 25%. The corresponding decreases are 10 times at low temperatures whereas they reach up to 1000 times for the high temperatures. Since reduced dielectric loss is desirable for capacitor applications, the composites having high KS additive % may have the potential to be utilized in energy storage implementations.

4. DISCUSSION

To evaluate all the results given in Sections 3.3.1 and 3.3.2 together and to specify the usability of the composites in terms of energy applications, ε' and ε'' have been investigated depending on the temperature and KS contribution at two spot frequencies, one corresponding to the low and one for the high frequency. The variations of ε' and ε'' with temperature and reinforcing additive percentage have been shown for 100 Hz and 0.1 MHz in Figure 10.

Figure 10. *The variations of the real and imaginary parts of permittivity with frequency and temperature for a,c) 100 Hz and b,d) 0.1 MHz spot frequencies*

As shown in Figure 10(a) and (c), while the ε' values at room temperature are almost the same at 100 Hz for all KS additive percentages, the high ε' values can be achieved between 12.5% and 20% KS contribution at high temperatures which are above 340K. When the working frequency reaches 0.1 MHz, the higher ε' values have been determined for the KS addition percentage interval of 15-22.5% at all temperatures. Moreover, the real part of the complex permittivity had an increasing trend with increasing temperature. Furthermore, KS additives above 25% can be considered as the saturation contribution level for PPy since these additive percentages create a decrease in ε' compared to pure PPy.

The energy loss of the samples can also be discussed in the context of the low and high frequencies. As is seen from Figure 10(b) and (d), the high dielectric loss values have been observed as the temperature increases from room temperature to 353 K. Additionally, the highest ε'' values have been determined for the KS additive interval of 12.5-22.5% at high temperatures for both frequencies. The lowest dielectric losses at all temperatures have been figured out both in the vicinity of the additive level of 7.5% and the addition of more than 25% KS.

5. CONCLUSIONS

In the present study, different ML models have been implemented for the estimation of the dielectric performance of the PPy based KS reinforced composites. After performing LR, SVM, and GPR algorithms, it has been revealed that among these models, GPR is the best suitable algorithm to estimate the frequency and temperature-dependent dielectric permittivity with the lowest error. Based on the good agreement between experimental data and GPR model predictions, initially, the GPR model has been used to predict the ε' and ε'' parameters of the samples for experimentally untreated temperatures including 303, 323, and 343 K. Then, the predictions of the ε' and ε'' for the new PPy based composites consisting of 7.5%, 15%, 25%, and 30% KS additives which were not produced experimentally for the temperature interval of 296-353 K have been performed.

The results of this study can be evaluated in terms of both the energy storage performance of PPy and machine learning researches.

In the light of both experimental and GPR estimations, the performance of composites for energy storage applications can be summarized as follows:

- ✓ Although the 7.5% KS additive did not cause a significant increase in the real part of the complex permittivity of PPY, since it reduced the dielectric loss significantly, PPy/7.5% KS composite can be utilized in energy storage applications.
- ✓ Due to higher ε' and relatively lower ε'' values obtained in the low and high-frequency range over the entire temperature band in comparison to pure PPy for the 20% -22.5% KS additive range, it can be suggested that this KS additive interval makes PPy more suitable dielectric layer for capacitor applications.
- ✓ 15% -17.5% KS additive level increased ε' and ε'' values at both the low and high frequencies and this increase continued as the temperature increases. In this context, the composites consisting of KS between 15% and 17. 5% can be used in decoupling capacitors.
- ✓ Finally, although the ε' value is reduced by 10 times at low frequencies for the addition of KS in 30%, since the dielectric loss is reduced by almost 10^4 times, PPy/30% KS composite can be preferred for energy storage applications.

On the other hand, this study primarily may set an example for new studies for dielectric parameter determination with ML. Moreover, in the future, it can work on an algorithm that can simultaneously estimate the ε' and ε'' parameters as a function of frequency, temperature, and additive concentration rate. If this can be achieved, it may be possible to keep production and characterization costs to a minimum by first determining the suitable working temperature, operating frequency range, and contribution percentage with ML for the production of dielectric material with desired properties.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Raimbault, N., Grisafi, A., Ceriotti, M., Rossi, M., "Using Gaussian Process Regression to Simulate the Vibrational Raman Spectra of Molecular Crystals", New Journal of Physics, 21: 105001, (2019).

[2] Pilania, G., Balachandran, P., Gubernatis, J. E., Lookman, T., "Classification of ABO3 perovskite solids: a machine learning study", Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials, 71: 507-513, (2015).

[3] Pilania, G., Gubernatis, J. E., Lookman, T., "Structure classification and melting temperature prediction in octet AB solids via machine learning", Physical Review B, 91: 214302, (2015).

[4] Huan, T.D., Mannodi-Kanakkithodi, A., Ramprasad, R., "Accelerated materials property predictions and design using motif-based fingerprints", Physical Review B, 92: 014106, (2015).

[5] Rupp, M., Tkatchenko, A., Müller, K.-R., Von Lilienfeld, O.A., "Fast and accurate modeling of molecular atomization energies with machine learning", Physical Review Letters, 108: 058301, (2012).

[6] Dey, P., Bible, J., Datta, S., Broderick, S., Jasinski, J., Sunkara, M., Menon, M., Rajan, K., "Informatics-aided bandgap engineering for solar materials", Computational Materials Science, 83: 185-195, (2014).

[7] Lee, J., Seko, A., Shitara, K., Nakayama, K., Tanaka, I., "Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques", Physical Review B, 93: 115104, (2016).

[8] Pilania, G., Wang, C., Jiang, X., Rajasekaran, S., Ramprasad, R., "Accelerating materials property predictions using machine learning", Scientific Reports, 3: 2810, (2013).

[9] Guo, D., Wang, Y., Nan, C., Li, L., Xia, J., "Application of artificial neural network technique to the formulation design of dielectric ceramics, Sensors and Actuators A: Physical, 102: 93-98, (2002).

[10] Habibi-Yangjeh, A., "Prediction dielectric constant of different ternary liquid mixtures at various temperatures and compositions using artificial neural networks", Physics and Chemistry of Liquids, 45: 471-478, (2007).

[11] Kılıç, M., Eyecioğlu, Ö., Özdemir, Z.G., Alkan, Ü., "DYPE/PANI kompozit filmlerin sıcaklığa ve PANI katkı konsantrasyonuna bağlı olarak dielektrik parametrelerinin GRSA ile tahmini", Journal of The Faculty of Engineering and Architecture of Gazi University, 35: 1077-1088, (2020).

[12] Mannodi-Kanakkithodi, A., Pilania, G., Huan, T.D., Lookman, T., Ramprasad, R., "Machine learning strategy for accelerated design of polymer dielectrics", Scientific Reports, 6: 20952, (2016).

[13] Schütt, K. T., Glawe, H., Brockherde, F., Sanna, A., Müller, K.-R., Gross, E.K., "How to represent crystal structures for machine learning: Towards fast prediction of electronic properties", Physical Review B, 89: 205118, (2014).

[14] Schmidt, E., Fowler, A. T., Elliott, J. A., Bristowe, P. D., "Learning models for electron densities with Bayesian regression", Computational Materials Science, 149: 250-258, (2018).

[15] Owolabi, T. O., Akande, K. O., Olatunji, S. O., "Prediction of superconducting transition temperatures for Fe- based superconductors using support vector machine", Advances in Physics Theories and Applications, 35: 12-26, (2014).

[16] Ponte, P., Melko, R. G., "Kernel methods for interpretable machine learning of order parameters", Physical Review B, 96: 205146, (2017).

[17] Caro-Gutiérrez, J., González-Navarro, F. F., Curiel-Álvarez, M. A., Peréz-Landeros, O. M., Valdez-Salas, B., Radnev-Nedev, N., "Machine learning for predicting the average length of vertically aligned TiO_2 nanotubes", AIP Advances, 10: 075116, (2020).

[18] Zhang, Y., Xu, X., "Curie temperature modeling of magnetocaloric lanthanum manganites using Gaussian process regression", Journal of Magnetism and Magnetic Materials, 512: 166998 (2020).

[19] Yang, K., Huang, X., Huang, Y., Xie, L., Jian, P., "Fluoro-polymer@ $BaTiO_3$ hybrid nanoparticles prepared via RAFT polymerization: toward ferroelectric polymer nanocomposites with high dielectric constant and low dielectric loss for energy storage application", Chemistry of Materials, 25: 2327-2338, (2013).

[20] Bishop, C. M., Pattern Recognition and Machine Learning, Springer, USA, (2006).

[21] Rasmussen, C. E., Williams, C. K., Gaussian Processes for Machine Learning, The MIT Press, USA, (2006).

[22] Rasmussen, C. E., Gaussian processes in machine learning, in: Summer School on Machine Learning, Springer, (2003).

[23]   Neal, R. M., Bayesian learning for neural networks, Springer Science & Business Media, New York, (2012).

[24]   Kılıç, M., "Natural additive material for desirable dielectric properties of polypyrrole: Limestone", Synthetic Metals, 260: 116297, (2020).