Chen *et al. Energy Mater.* **2025**, *5*, 500120
**DOI:** 10.20517/energymater.2024.311

**Energy Materials**

**Review**

**Open Access**

# Statistical and artificial intelligence approaches towards the optimization of thermoelectric materials synthesis: a review

**Wei-Hsin Chen**[1,2,3,*], **K. Aishwarya**[4], **Kripasindhu Sardar**[1]

[1]Department of Aeronautics and Astronautics, National Cheng Kung University, Tainan 701, Taiwan.
[2]Department of Chemical and Materials Engineering, Tunghai University, Taichung 407, Taiwan.
[3]Department of Mechanical Engineering, National Chin-Yi University of Technology, Taichung 411, Taiwan.
[4]Department of Physics, Vellore Institute of Technology, Chennai 600127, India.

***Correspondence to:** Dr. Wei-Hsin Chen, Department of Aeronautics and Astronautics, National Cheng Kung University, No. 1, University Rd., Tainan 701, Taiwan. E-mail: chenwh@mail.ncku.edu.tw

**How to cite this article:** Chen, W. H.; Aishwarya, K.; Sardar, K. Statistical and artificial intelligence approaches towards the optimization of thermoelectric materials synthesis: a review. *Energy Mater.* **2025**, *5*, 500120. https://dx.doi.org/10.20517/energymater.2024.311

**Received:** 31 Dec 2024 **First Decision:** 24 Feb 2025 **Revised:** 13 Apr 2025 **Accepted:** 27 Apr 2025 **Published:** 13 Jun 2025

**Academic Editor:** Yuhui Chen **Copy Editor:** Fangling Lan **Production Editor:** Fangling Lan

## Abstract

Thermoelectric (TE) materials, capable of directly converting heat to electricity, offer a promising sustainable energy and waste heat recovery solution. Despite extensive research, a significant bottleneck remains: the synthesis of high-performance TE materials still relies heavily on trial-and-error approaches, which are time-consuming and resource-intensive. Moreover, while machine learning (ML) and design of experiments (DOE) have shown potential in optimizing synthesis processes across materials science, their systematic application to TE materials remains underexplored. In particular, very few reviews have addressed the integration of statistical and AI-guided methods for synthesizing and optimizing TE materials. This manuscript comprehensively reviews recent advances in statistical and artificial intelligence techniques for optimizing TE material synthesis. It first discusses the role of DOE in identifying critical synthesis parameters and explores various ML methods for predicting TE performance. This study then highlights case studies involving different TE material systems, synthesis strategies (e.g., ball milling, sputtering, electrodeposition), and ML-based performance prediction and optimization. This work fills a critical gap by linking data-driven optimization techniques with experimental synthesis in the TE field. It not only consolidates current knowledge but also sets the stage for future studies aiming to bridge material discovery and practical manufacturing. The insights presented are instrumental in accelerating the development of next-generation TE devices.

**Keywords:** Thermoelectric generator (TEG), materials, synthesis, optimization, design of experiment, machine learning

## INTRODUCTION

The ever-growing energy demand has become a global challenge due to the limited reserve of fossil fuels. There have been considerable efforts to develop alternatives to fossil fuels, such as solar, nuclear, and wind. Naturally, a lot of energy is wasted as heat in the modern world. Thermoelectric (TE) technologies, which utilize the phenomena of energy conversion between heat and electricity, have the potential to reuse waste heat for sustainable electrical energy generation[1]. Nevertheless, TE technologies are environmentally friendly renewable energy sources and have drawn tremendous attention over the last few decades[2].

Due to the ability to convert thermal to electrical energy and vice-versa, TE materials have potential in various applications, including refrigeration, power generation, and waste heat recovery[3]. Recently, there have been plenty of research activities in the area of the use of statistical methods such as design-of-experiment (DOE), machine learning (ML), and artificial intelligence (AI) in the area of TE technology with a focus on the design of TE materials and thermoelectric generators (TEGs).

DOE is a statistical way of optimizing the response of experiments. Constructing experiments with the minimum optimization parameter variations statistically to enhance the system's performance is crucial. AI is a technology operating to do the task, which is constructed through data collection, model training, optimization and deployment. It is used to find the new materials in the TE. ML belongs to AI, a statistical model for optimization and finding the relation between input and output parameters. ML is developed with mathematical models and algorithms for finding the pattern through statistical analysis of input data. Gorai *et al.* have reviewed optimizing materials properties and designing and discovering TE materials using ML[4]. Wang *et al.* have given an overview of the use of several ML methods, such as Bayesian optimization, regression, and neural network (NN) models in TE research[5]. Chen *et al.* have reviewed the use of ML in discovering and designing various materials for energy-related applications. This included ML in photovoltaics, batteries, catalysis, and thermoelectric[6]. Recatala-Gomez *et al.* reviewed the historical evolution of various inorganic TE materials. They postulated that combining data generation, ML, high-throughput synthesis and characterization, and high-performance computing can accelerate the discovery of novel TE materials[7]. Recently, Wang *et al.* have critically reviewed the progress on the application of ML in (i) predicting and optimizing the properties (electrical and thermal transport) of TE materials; and (ii) the designing and screening of TE materials[8]. Furthermore, the optimization of TEGs based on statistical approaches such as the Taguchi method, the response surface methodology (RSM), and the analysis of variance (ANOVA) has been reviewed by Chen *et al.*[9]. Kucova *et al.* reviewed waste heat harvesting from the Internet of Things (IoT) through ML[10]. The review concluded the suitability of TEG in low-grade waste heat harvesting through the results of various ML algorithms. Song *et al.* draw the roadmap from high throughput materials discovery to advanced device fabrication[11]. Also, this review discussed the discovery of new TE material through ML algorithms. Deng *et al.* discussed predictive ML algorithms based on the TE materials[12]. The previously reported reviews on TE mainly focused on new material discovery based on the ML algorithm. One important part that was not discussed is atomic characteristics relation with TE parameters and the parameter optimization in material synthesis through ML approach. However, these are important to understand the TE material tailored to find a suitable approach to tune the TE performance with minimal optimization parameters.

Regardless of the ability of ML and DOE to direct the discovery and optimization of TE materials and TEGs, the ultimate ability to control the synthesis in a rational and controllable way will determine the research and development of future TE technologies. However, the synthesis attempt of existing and unknown TE

materials from existing literature based on knowledge of materials synthesis and the prediction and control of outcome is complicated. This is because the control over TE materials synthesis depends on carefully controlling many experimental parameters such as choice of precursors, synthetic methods, temperature, pressure, atmosphere, time, additives, *etc*.

There are a total of 38,413 published articles under the keyword 'thermoelectric materials' over the past ten years in the Science Direct database. Figure 1 shows how publication activity in TE materials has changed over the last ten years. The surge in TE materials research is apparent, and the number of publications has increased rapidly. Therefore, the research areas involving TE materials are popular among the scientific community. However, controllable synthesis of TE materials is a significant bottleneck for realizing practical thermoelectric devices such as TEGs. Notably, there is a knowledge gap between implementing statistical optimization methodologies, discovering new TE materials, and finding optimal synthesis and device fabrication conditions. This article covered the importance of ML in DOE for TE material synthesis, material chemistry parameters involving boosting TE performance, segregating materials based on performance, finding new material, and optimizing composition. In TE material creation such as a thin film, the ML approach in parameter optimization for coating and sintering is discussed.

## PERFORMANCE ENHANCEMENT OF THERMOELECTRIC MATERIALS

The efficiency of TEGs depends on the performance of TE materials. In this direction, significant advancement has been achieved through research and development toward the optimal performance of TE materials[13-19].

The performance of TE materials is usually expressed as the "figure of merit" $ZT = S^2\sigma T/ \kappa_e + \kappa_l$, where the Seebeck coefficient is $S$, the electrical conductivity is $\sigma$, and electronic and lattice thermal conductivities are $\kappa_e$ and $\kappa_l$ espectively, at temperature $T$. Here $(\kappa_e + \kappa_l)$ is known as the total thermal conductivity of TE materials. Thus, a high value of ZT implies high performance of TE materials. Moreover, TE materials with high ZT values can lead to efficient TE devices for commercial applications. The high value of ZT at a given temperature can be achieved with TE materials having high S, high $\sigma$, and low thermal conductivities $\kappa_e$ and $\kappa_l$. However, $\sigma$ decreases when $S$ increases. On the other hand, total thermal conductivity $(\kappa_e + \kappa_l)$ of TE materials is proportional to $\sigma$. Due to this complex interrelation between the TE parameters, it is pretty challenging to control independent parameters experimentally for the optimization of the performance of TE materials. Literature-reported methods for optimizing ZT focused on the methods for reducing thermal conductivity and enhancing $S$ and $\sigma$, respectively[17,20]. Finally, when the thermal conductivity of materials is very low or unavailable, the performance of TE materials can be expressed as the power factor $S^2\sigma$.

### Thermoelectric materials development

The range of materials that exhibit TE properties encompasses wide varieties, including inorganic, organic, and inorganic-organic hybrid[21]. Naturally, each type of material has its advantages and disadvantages. There are vast numbers of inorganic materials that exhibit TE properties. Typical inorganic TE materials are based on chalcogenides (such as $Bi_2Te_3$, PbTe, SnSe, *etc*.), Si-Ge alloys, multicomponent oxides, skutterudite-type materials, half-Heusler alloys, and clathrates[14,19]. The variation of ZT with temperature of some typical inorganic TE materials is shown in Figure 2[22]. As a result of extensive research and development over the past few decades, $Bi_2Te_3$ and Si-Ge-based materials are used in commercial TE devices. The main advantage of inorganic TE materials is their high TE performance. Commercial TE devices use inorganic materials with ZT~1. Therefore, searching for new materials with high ZT values has been popular in TE material research[20,23]. The disadvantages of inorganic TE materials include poor abundance of elements, high cost, and toxicity[24]. On the other hand, the advantages of organic TE materials

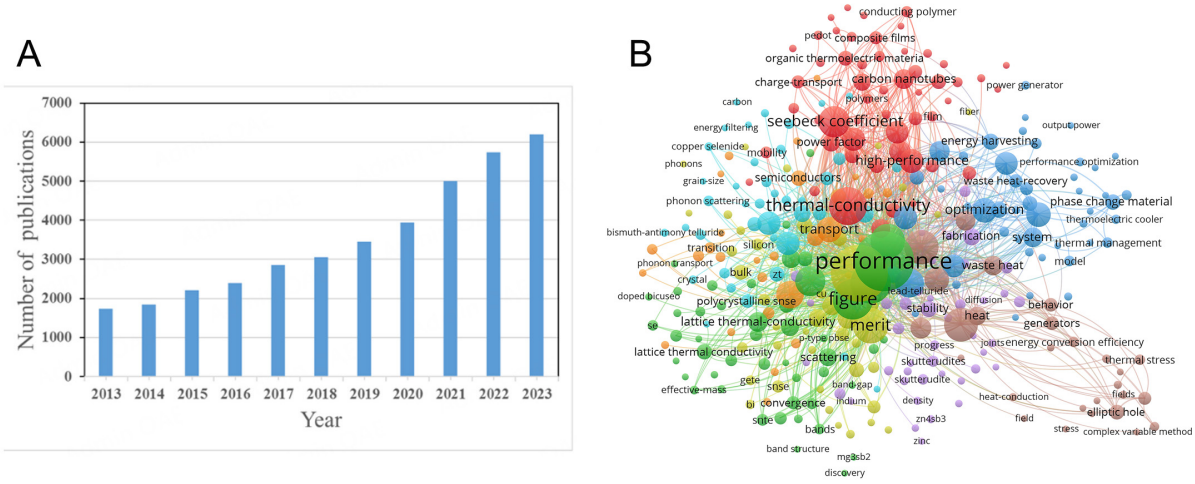


**Figure 1.** (A) Showing histogram of several results as a function of publication years with keywords "thermoelectric materials" in the Science Direct database. (B) Map created by VOSViewer for "thermoelectric materials" from Web of Science.

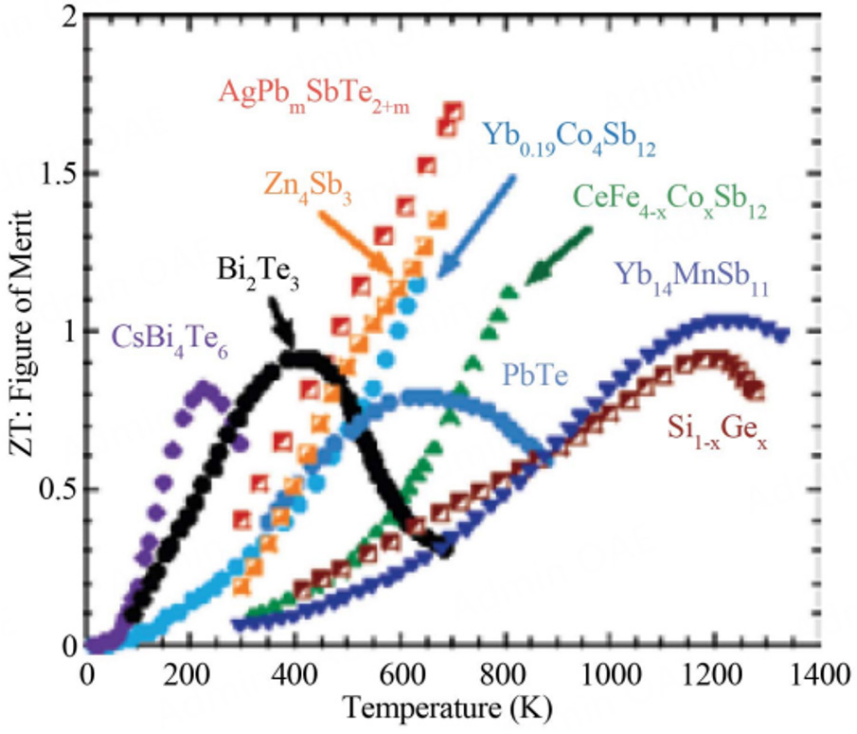


**Figure 2.** Variation of ZT with temperature of some common inorganic TE materials. Adapted from Jarman *et al.*[22]. (Licence CC By 4.0).

include earth-abundant elements, ease of manufacturing involving solution-based methods, the possibility of a cost-effective scale-up process, lightweight, mechanical flexibility, low toxicity, and low thermal conductivity[25,26]. However, due to poor Seebeck coefficient and electrical conductivity, organic materials have inferior ZT values[27]. Nevertheless, the development of organic materials for TE devices has been reviewed by several authors[25,27-29]. Interestingly, the inorganic-organic hybrid materials are emerging as novel TE materials. The hybrid TE materials could combine the high *S* and σ of inorganic materials and the

low thermal conductivity of organic materials[21,30].

### Design of experiment for TE materials synthesis

During the optimization of synthesis of TE materials, experiments are performed to measure the effects of experimental variables on responses. The optimization of preparation involves finding a combination of variables that gives the best response. Thus, there are many types of experiments. DOE is a statistical way of optimizing the response of experiments. An efficient DOE is a sequential approach wherein information gained in the initial stage can be used to decide which factors may be kept constant or varied in later stages. In this process, reducing the number of experimental variables to a more significant one is possible. Thus, statistical DOE through fractional factorial design can provide an opportunity to minimize the number of experiments by optimizing experimental conditions. The main advantages of DOE are efficiency, response analysis, and simple interpretation of statistical outcomes.

The DOE has existed for quite some time, but its use is very new in synthesizing TE materials. Selvaratnam *et al.* have reviewed the application of ML in general materials chemistry[31]. Recent progress in DOE and ML for predicting and controlling inorganic materials synthesis has been discussed by Willamson *et al.*[32]. Recent review articles on the use of DOE, ML, and AI for the synthesis of general solid-state materials have been presented by Baum *et al.*[33].

The first example of the application of statistical DOE in TE materials synthesis is for the optimization of the sintering parameters of $K_2Bi_8Se_{13-x}S_x$ (for x = 0, 4, 6, and 8)-based TE materials by Kyratsi *et al.*[34]. Sintering conditions must be optimized to obtain high-quality pellets of TE materials with the highest density. In this work statistical DOE through fractional factorial experiment and Taguchi table has been employed for identifying optimum parameters for sintering (duration, temperature, and pressure) for the fabrication of high-quality pellets. Based on the ANOVA, it has been postulated that the density of the pellet is more significantly affected by pressure than duration and temperature. The best condition for hot-press sintering of these TE materials has been 80 MPa pressure, 530 °C temperature, and 90 min duration, respectively, with a relative density of pellet approaching ~97%. These optimized sintering conditions have been used for sintering the pellets of the entire $K_2Bi_8Se_{13-x}S_x$ compounds. Interestingly, the sample with x = 0 exhibited a ZT value of 0.58 at 673 K.

$Bi_2Te_3$ and its alloys, such as $Bi_{2-x}Sb_xTe_3$ and $Bi_2Te_{3-x}Se_x$, have been used in TE technologies for decades. However, Kanatzia *et al.* have employed factorial DOE via ANOVA for the ball-milling synthesis of $Bi_2Te_3$ for the first time[35]. The analysis of parameters for ball milling, such as duration, rotation speed, and ball-to-material ratio, respectively, suggests a strong influence of the duration and speed of ball milling on the TE properties of nanocrystalline $Bi_2Te_3$. The ZT value of optimized ball-milled nanocrystalline $Bi_2Te_3$ is 0.72 at 100 °C. It is expected that DOE can further improve the ZT value of $Bi_2Te_3$-based TE materials through efficient optimization of ball-milling parameters.

Nuthongkum *et al.* applied RSM and statistical central composite design (CCD) for modeling and optimizing thin film deposition parameters of thermoelectric $Bi_2Te_3$ material using the radiofrequency (RF) magnetron sputtering technique[36]. This study has statistically analyzed the formation of Bi-Te thin films considering several deposition factors, such as the flow rate of Ar gas and annealing temperature. The DOE using RSM to find the optimal conditions for the targeted response utilizes mathematical and statistical methods to develop models and evaluate factors. Thus, RSM plots suggest that the concentration of Te in Bi-Te thin film would decrease if the annealing temperature is increased at a lower Ar flow rate. Using this model, it has been determined that the optimized conditions for the deposition of good quality thin films

$Bi_2Te_3$ using RF magnetron sputtering technique are 285 °C temperature of annealing and 103.5 sccm flow rate of Ar. However, the TE properties of stoichiometric $Bi_2Te_3$ films using this press have been poor, and further optimization of additional deposition parameters (for example, annealing and substrate temperatures) has been proposed. Furthermore, in the study of Khumtong *et al.*, the RSM based on CCD has been employed to optimize thin films of $Sb_2Te_3$ deposition using the RF magnetron sputtering technique[37]. It has been found that the highest TE power factor of $Sb_2Te_3$ thin films requires deposition parameters Ar gas flow rate 120 sccm and processing temperature 375 °C. The highest power factor value obtained for stoichiometric $Sb_2Te_3$ film is $2.0 \times 10^{-3}$ W/mK$^2$ at 250 °C.

Carbon materials are widely studied in electronic devices due to their availability in different structures[38]. The high electrical conductivity is crucial for TE performance enhancement. Jagadish *et al.* developed a DOE study on the investigation of TE properties of thin films of $Bi_2Te_3$ on recycled carbon fiber by electrodeposition technique[39]. Herein, the DOE based on the D-optimal model under RSM has been adopted to optimize the combined effect of electrodeposition parameters, leading to the optimum Seebeck coefficient. The authors used a multivariate approach involving DOE based on 23 runs to optimize experimental parameters. The optimum electrodeposition parameters are potential -0.10 V, time 0.5 h, temperature 25 °C, and electrolyte compositions 0.240 Bi/(Bi + Te). Thus, using DOE, the optimized electrodeposition parameters yielded an experimental Seebeck coefficient of -13.42 μV/K, about 33% larger than samples prepared without DOE.

Lead chalcogenides (PbTe, PbSe, PbS) are excellent TE materials. In particular, $Pb_{1-x}Sn_xTe$-based materials are promising for TE applications due to their high ZT values. Recently, Sam *et al.* have developed a DOE method for synthesizing $Pb_{1-x}Sn_xTe$ (x = 0.67) crystals by horizontal vapor phase growth method[40]. This study employs $2^k$ full factorial design to correlate ZT with the effect of synthesis temperature and time. Based on the ANOVA of ZT, the synthesis temperature and time exhibit significant ZT response. Thus, the crystals prepared at 1,200 °C for 4 h have exhibited the highest ZT of approximately 0.084.

Due to high operating temperatures, SiGe alloy-based TE materials are widely used in radioisotope TEGs[41]. The n-type SiGe alloy may exhibit a very high ZT value of 1.84 at 1,073 K[42]. Ahmad *et al.* studied statistical DOE and RSM optimization of ball milling and hot press sintering parameters to prepare SiGe alloy[43]. In this study for TE properties of SiGe alloy, 13 experiments were incorporated in two-factor CCD for obtaining regression equations. For the maximum ZT, the hot press temperature is 1,504.5 K, and the time for ball milling is 53.6 h. Herein, the predicted and experimental optimum ZT values were 1.148 and 1.146, respectively.

Effective heating and cooling on the surface of TEG increases the temperature difference across TEG, which directly relates to device efficiency and output power. However, the construction of the microstructure of TE materials is vital to withstand significant temperature differences for performance enhancement. The optimization of TE legs (such as the areas of hot and cold sides, pin configuration, contact angle, *etc.*) and the construction of the 3D architecture of TE legs tailors to avoid the brittleness of material and the heat stagnation in its legs[44]. Maduabuchi[45] studied thermomechanical optimization using a deep neural network (DNN). The data obtained from the finite element method was fed to the DNN to learn the pattern. The well-fitting of experimental and DNN results indicated the suitability of DNN, reducing the simulation time. The temperature difference between the hot and cold sides was crucial to maximize the output power. Besides, this report discussed the optimization of temperature effect on temperature on the hot side ($T_h$) and temperature on the cold side ($T_c$) by the wind speed. This effective cooling was achieved by installing low thermal conductivity material on the cold side, and the cold side temperature was

optimized with a cooling coefficient. Besides, the increment in $T_h$ increased the thermal stress on the legs, which accounted for performance reduction. Meanwhile, the optimized geometrical parameter of TEG minimized the thermal stress, avoiding the device's cracks and heat stagging. In the staggered and integrated device, mechanical stability was a crucial parameter. The same author reported the optimization focused on thermomechanical properties in the TEG and solar cell integrated device. The numerical method generated data was used to feed the DNN to predict the performance quickly and with more accuracy. After the optimization by focusing on the thermo-mechanical stability, the output power was increased, and the thermal stress was reduced[46].

Designing an ML model based on actual environmental conditions is crucial, which helps design the TEG for large-scale applications and suitable environmental conditions. However, the temperature of the TEG surface depends on the environmental conditions. Ameenuddin Irfan *et al.* designed the TEG based on the actual environment by analyzing the humidity variation and actual room temperature, predicting $T_h$ and $T_c$[47]. For this, 35-day real-time data was recorded, which was used to feed the ML model to analyze the variation. ML models, such as linear, tree regression, and Gaussian process regression (GPR), were employed, and the GPR model had a high accuracy in predicting the environmental conditions.

## MACHINE LEARNING (ML) METHODS

In ML methods, problems are solved by developing mathematical models and algorithms by discovering patterns through statistical analysis of input data. The most common ML algorithms in chemistry are support vector machine (SVM), random forest (RF), k-nearest neighbors (kNN), ensemble, decision tree (DT), and NN methods, respectively.

### Machine learning in materials chemistry

In materials chemistry, AI technologies, including ML methods, have been increasingly used to predict the crucial factors for the synthesis and select the optimal reaction conditions. Gulevich *et al.* have reviewed the application of ML methods to develop synthesis and choose the best synthesis conditions for colloidal nanomaterials[48]. As shown in Figure 3, AI technologies have several benefits for synthesizing colloidal nanomaterials[48]. Thus, optimization of synthesis parameters (time, temperature, concentration of precursors, and additives) necessary for the synthesis of nanocrystals of several chalcogenides (CdS, CdSe, PbS, and ZnSe) have been studied using ML[33]. The use of similar ML methods for synthesizing size and shape-controlled chalcogenides-based TE nanomaterials is yet to be known. This is one area that will expand in the coming future.

The relation between the structural parameters in chemistry and the thermoelectric parameters is crucial to segregating the material from high performance to low performance. Compared with the conventional cascaded arrangement of TE devices, low-cost spin-driven thermoelectric (STE) consisting of simple layered structures gained much attention since it has been fabricated by sputtering, coating, and platting, which is a direct approach. However, STE consists of rare earth elements that are pretty costly. Applying ML to analyze suitable composition is more time-efficient and can reduce material waste.

Iwasaki *et al.* applied the ML approach to discover novel STE material through actual material synthesis[49]. The experimental data of series of rare earth substituted yttrium iron garnet Pt/R:YIG (R = La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, and Lu, except Pm) fabricated on $Gd_3Ga_5O_{12}$ (GGG) and $Gd_{2.675}Ca_{0.325}Ga_{4.025}Mg_{0.325}Zr_{0.65}O_{12}$ (SGGG) substrates taken to analyze the Seebeck coefficient ($S_{STE}$) relation with other parameters such as atomic weight $n_R$, spin and orbital angular momenta $S_R$ and $L_R$, lattice mismatch $\Delta a$ between R:YIG and the substrate under spin driven effect. In this study, four types of ML

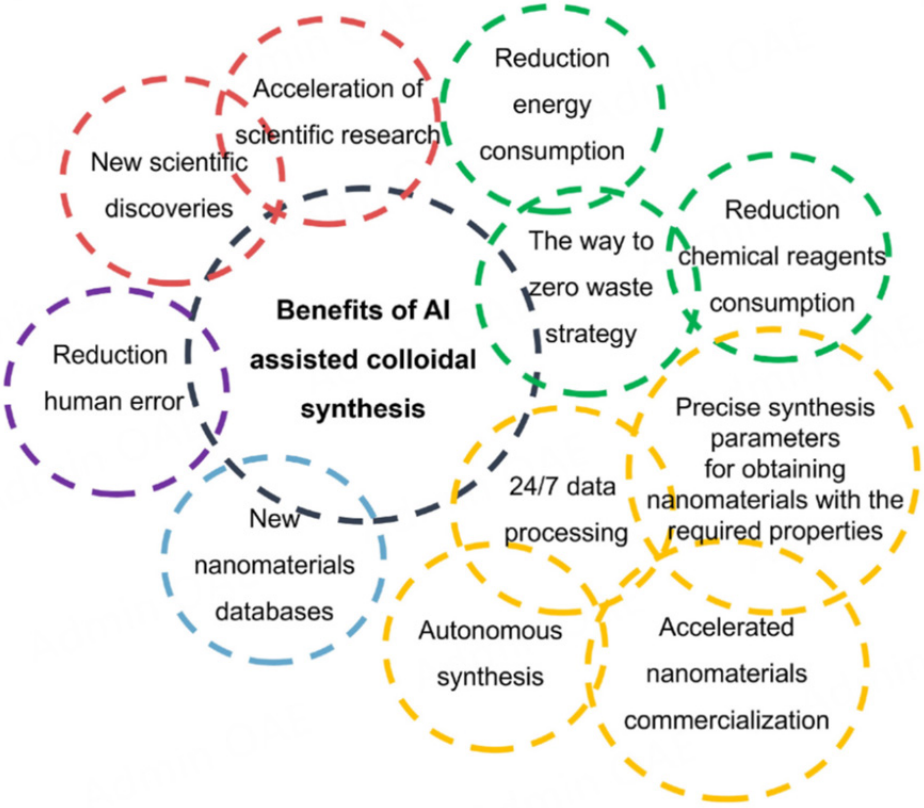


**Figure 3.** Showing the benefits of using AI for synthesizing colloidal nanomaterials[48]. (License Number 6015211273249).

models were applied: decision tree regression (DTR), elastic net (EN), quadratic polynomial-least absolute shrinkage and selection operator (QP-LASSO), and NN to predict the relationship between the parameters. Thus, four ML algorithms converge, with $S_{STE}$ showing a positive connection with $n_R$ and $L_R$ but a negative correlation with $\Delta a$ and $S_R$. The positive correlation of $L_R$ with thermopower tested with another material Fe-Pt-T (T = Sm, Gd, Cu, and W) ternary alloys, relying on anomalous Nernst effect (ANE) originating in the spin-orbit interaction. $L_R$ value for Sm was larger than Gd, Cu, and W, while this material has $L_R = 0$. The composition around $Fe_{0.665}Pt_{0.27}Sm_{0.065}$ shows a large $S_{STE}$ at least one order greater than the other ANE materials. ML is a suitable approach to finding suitable TE parameters among clusters of materials (a combination of various periodic table elements) in relation to their crystal structure, compound chemistry, and interatomic bonding. Finally, the linear fitting of predicted $S_{STE}$ through ML and the experimental $S_{STE}$ indicates better accuracy. Among four MLs, NN gives better accuracy, which means well-matched experimental and predicted values.

Further, analyzing new materials and their TE properties is crucial. Tewari *et al.* applied the ML approach to classify thermal conductivity among oxides and oxide alloys of transition metals, that is, elements of groups 3-11 and periods 4-6[50]. A two-step ML model was employed: one is classification, and another is regression. The complete set-off ML model is shown in Figure 4. Gradient boosted tree classifier was used to predict the key material properties influencing $\kappa_l$ and to classify material with low, medium, and high $\kappa_l$. These key material properties were lattice energy per atom, atom density, band gap, mass density, and oxygen ratio by transition metal atoms. Above mentioned properties define the crystal structure, compound chemistry, and interatomic bonding of a compound. The regression model was employed to predict absolute $\kappa_l$ with various models, including Cubist, GPR polykernel, kNN, GBM, XGBoost, RF, Kernel ridge, and deep neural nets.

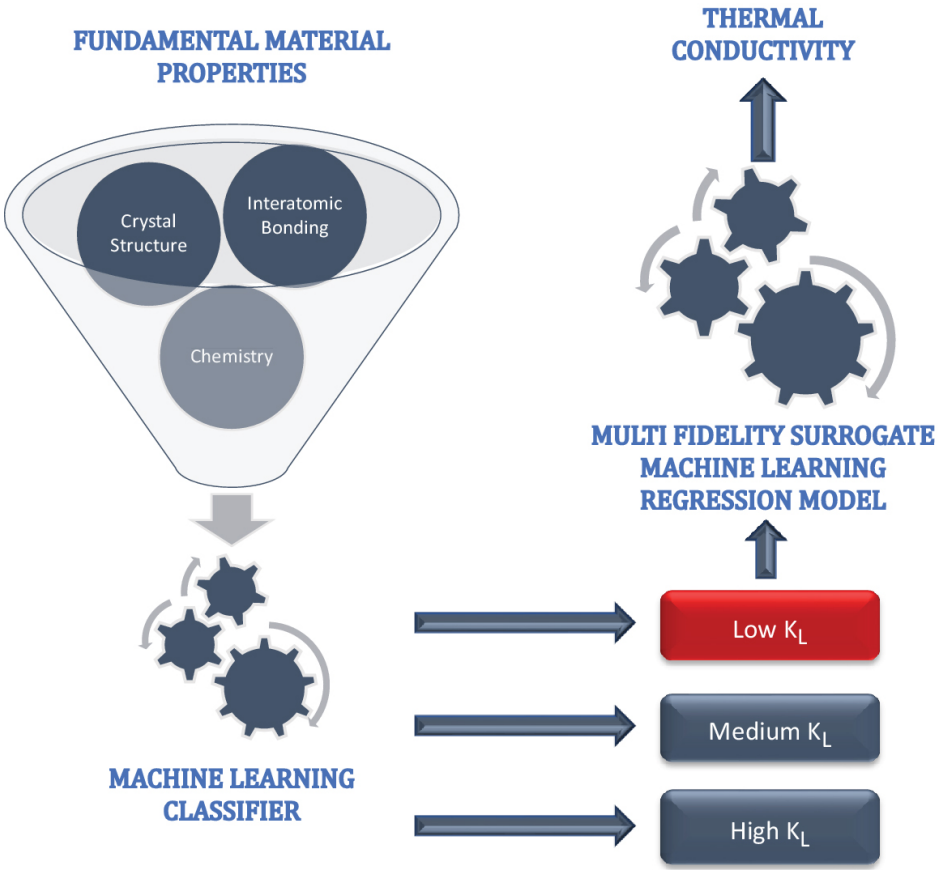


**Figure 4.** ML approach to predict low $\kappa_l$ value among oxide-based alloys[50]. (License CC By 4.0).

Among these, Cubist, GPR polykernel, GBM, and Kernel ridge models have better fit with actual value with the coefficient of determination ($R^2$) > 0.9, which is mainly used on low $\kappa_l$ materials. Another report by Juneja *et al.* detailed the segregation of low and high $\kappa_l$ among 120 materials for TE and thermal barrier coatings, respectively[51]. These materials' maximum phonon frequency integrated Gruneisen parameter up to 3 THz, average atomic mass, and unit cell volume obtained from material project relay on $\kappa_l$. GPR was employed as an ML model to predict and segregate the materials based on $\kappa_l$. The data was collected from the material project, and the obtained log($\kappa_l$) values between experimental and ML approaches were analyzed with fitting. In this study, 15 new materials were identified from primitive and face-centered crystal classes, with a very low $\kappa_l$ range between 0.13 and 0.98 $Wm^{-1}K^{-1}$, considerable to TE performance. GPR model predicted the log-scaled $\kappa_l$ through ARD Matern 5/2 covariance function, which gave the train/test RMSE of 0.20/0.21 and the $R^2$ of 0.99/0.99. The same log-scaled $\kappa_l$ predicted through the slacker model overestimated the values and showed poor variability with the slacker model.

TE parameters are related to one another, raising the importance of finding the physical descriptors that connect all the TE parameters. Juneja *et al.* applied ML to predict the relation between elemental descriptors and electronic transport properties ($S$ and $\sigma$)[52]. Initially, 2838 compounds were screened based on the band gap (greater than zero) for nonmetallic compounds and the phonon frequency (greater than zero) for stability. After the screening process, 135 compounds remain and used for further process. GPR was developed as a predictor model by employing a 10-fold cross-validated least absolute shrinkage and selection operator (LASSO). Among 135 compounds, 34 compounds had a high-scaled power factor. In

them, 6 has $\kappa_l$ less than 2 $Wm^{-1}K^{-1}$. From these, a material $PdBr_2$ was identified with a high power factor and low lattice thermal conductivity. LASSO helped to screen the descriptor in relation to electronic transport parameters. The higher accuracies were obtained through the elemental and bonding descriptors. The transport parameters σ and *S* revealed the dependency relation with coordination number, boiling point, the heat of formation, and molar-specific heat coming under the bonding descriptor. From this, σ and *S* commonly relate to the bond strength. Further, $\kappa_l$ shows a close relation with the coordination number through anharmonicity. Volume is inversely correlated with the $\kappa_l$. Hence, electronegativity, bond strength, bond distance, volume, and coordination number have an excellent correlation with the electronic and thermal transport properties related to the chemical bonding descriptor. In conclusion, the chemical bonding-driven descriptor is a key point to connect the electronic and thermal transport properties. These reports find a common relation that could help to understand future experimental research in this domain. A total of 2,838 compounds were taken from different groups of materials: alkaline, transition, alkali, and lanthanide elements, and these were screened based on the band gap and phonon frequency; after that, 185 compounds were found to be nonmetallic and stable. Figure 5 illustrates the screening process, which includes the optimization, band gap, and phonon frequency. Also, another report has verified the relation between chemical bonding characteristics and $\kappa_l$. In this, the coordination number had a high value when the bond strength was weak and the bond distance was large[53].

Various atomic characteristics of materials that change the TE properties lead to identifying their relationship. Indeed, ML reduces the cost of material synthesis by predicting this relation. Li *et al*. applied the ML approach to analyze the TE performance of high entropy GeTe materials by correlating atomic features and the ZT[54]. Nine atomic features were selected for this analysis: atomic number, ionization energy, pseudopotential radius, atomic radius, electronegativity, electron affinity energy, molar volume, number of valence electrons, and atomic weight. Each element's atomic fraction is treated as a weighted score for its atomic properties. The following equation constructs the ML model in relation to atomic characteristics:

$$V_i = \sum v_{i,j} c_j$$

where the letters *i* and *j* represent certain elements, and $V_i$ is the average atomic characteristic, which is defined before ($V_1$ is the average atomic number, $V_2$ is the average atomic radius, $V_3$ is the average pseudopotential radius, $V_4$ is the average molar volume, $V_5$ is the average electronegativity, $V_6$ is the average ionization energy, $V_7$ is the average electron affinity energy, $V_8$ is average ionization energy, $V_9$ is the average atomic weight, $V_{10}$ is the average number of valence electrons, *T* is measuring temperature), $v_{i,j}$ represents the atomic characteristic, and $c_j$ is the atomic fraction. Different ML models are tested in this, namely LightGBM, surface vector regression (SVR), Ridge E, XGBoost, RF, and linear regression. LightGBM exhibits the least error value among these, with an $R^2$ of 0.954. Among the atomic characteristics, temperature has a greater impact on ZT, while $V_4$ and $V_5$ randomly decreased, and the optimum values of $V_4$ = 17.5 and $V_5$ = 2.05 were fixed in this. Then, the composition ratios of Ge, Te, Sn, and Mn are optimized, and the ZT values are tested at 790 K, the optimum temperature for this composition. Another composition set with Ge, Te, Sb, Sn, and Te was optimized, and the peak ZT is obtained at 800 K. Finally, the condition was applied practically, and the synthesized materials are new, with considerable ZT, ensuring low-cost synthesis.

The influence of input data size on training the ML model is vital to achieve accuracy. The sample data size defines the degree of freedom (DoF) on fitting. Zhang *et al*. studied the data size influence with descriptors

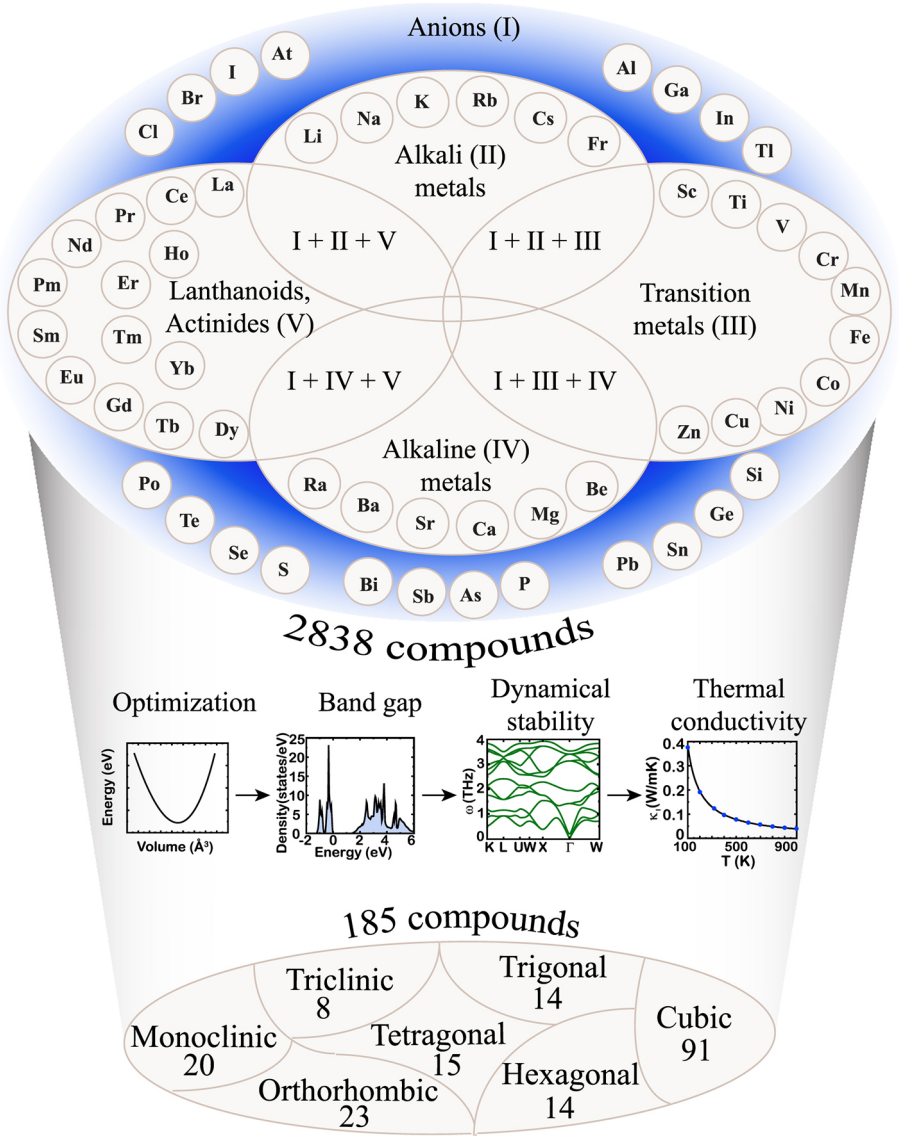


**Figure 5.** Process of screening compounds under the optimization, band gap, and phonon frequency to get nonmetallic and stable compounds, to calculate electronic properties[53]. (Creative Commons Attribution 4.0 license).

to predict the band gap, $\kappa_l$, and elastic properties of zeolites[55]. Rather than directly influencing model precision, the size of the dataset exerts its effect indirectly through the model DoF. It was challenging to predict the results in unknown domains without affecting the precision. Integration of crude estimation features in ML model improved the predictive accuracy without using the high DoF. Experimental and density functional theory (DFT)-based datasets were used to train the model before and after integrating the crude estimation, respectively. Kerne ridge regression ML model was used to predict the $\kappa_l$, the error of 6.2% that occurred was reduced to 4.1% after the integration of the crude estimation feature. This study implies the improvement of accuracy with the integration of crude estimation, which needs less data to create an ML model.

Doping is crucial for TE material performance enhancement. However, more elements can be used for doping in a particular site due to the availability of more elements in a similar oxidation state. Determining

the elements tailored to reduce material waste, cost, and time consumption is also crucial. He *et al.* reported the prediction of superior thermoelectric performance in unexplored doped-BiCuSeO via ML[56]. The experimental data of $Bi_{1-x}M_xCuSeO$ (M represents Ca, Mg, Sr, Ba, Na, La, Sm, Er, Ho, Cd, Sb, Pb, Ag, Al, Fe, K, Rb, Co, Cs, Mn, Nd, Sn, Yb, Zn, and Ni) was collected from the literature. The descriptors, a set of input parameters such as temperature, the content of the doping element, the relative molecular mass of $Bi_{1-x}M_xCuSeO$, the Mendeleev number of dopant, the Pauling electronegativity of dopant, the first ionization energy of dopant, the ionic radius of $Bi_{1-x}$ (r) and the Pauling electronegativity of $Bi_{1-x}$, where x ranges from 0 to 0.2 referred the doping content, used to generate a ML model. The workflow from data collection to ML prediction with the application of ML is shown in Figure 6A. Totally, six ML approaches were developed to solve this problem. Among these, GBR has an $R^2$ of 0.96, indicating better fitting with experimental input data. The correlation between descriptors and the ZT helps identify the new composition's ZT through the ML model. Figure 6B shows the ZT of the experimentally obtained value of doped elements and the predicted value. The optimized content of $Bi_{0.86}Po_{0.14}CuSeO$ and $Bi_{0.88}Cs_{0.12}CuSeO$ has improved ZT by 104% and 98% at 923 K, respectively. The employed ML approach identified how to analyze the suitable doping element among the more elements for TE performance enhancement with the input help of descriptors. Minhas *et al.* used the database of synthesized materials trained with ML models to predict the suitable doping elements to dope into GeTe, SnTe, PbTe, $Sn_{1-x}Se$, $Bi_{2-x}Sb_xTe_3$), skutterudite ($CoSb_3$, $As_2Te_3$), clathrates ($Ba_8Ga_{16}Ge_{30}$), and transition metal-based chalcogenides ($Cu_{2-x}Se$,$Ag_{2-x}Te_2$)[57]. Among various ML models, eXtreme gradient boosting regression (XGBR) has the best fitting with experimental database. The maximum ZT of 2.20 at 1,000 K was identified for $Bi_{0.1}Sb_{1.9}Te_3$. The correlation between the materials descriptors and the output parameter helped to find the new high-performance material. The descriptors selected were based on the material chemistry influencing the structural and transport properties. Also, the classification of the material was based on the performance of the RF classifier ML model in the different structures.

The ML model is accelerating to find a new TE material and highly efficient dopant. Parse *et al.* constructed an ML model to discover a new dopant to place in the Bi site in BiCuSeO[58]. The model design focused on the accuracy improvement by normalizing ZT of doped BiCuSeO with pristine BiCuSeO. The developed model produced the best fitting data with experimental. The $R^2$ value of 0.93 from the developed extra tree regression model was much closer to 1, indicating more accuracy than the initial model. New dopants were discovered through this model without wasting the material, which satisfied the thermoelectric principle. Based on various descriptors, the selected Si as a dopant for the Bi site improves the ZT by increasing mobility. Table 1 lists the survey of this review and refers to the descriptors, targets, and findings of the study related to ML in material chemistry.

### Machine learning in TE materials synthesis

The efficient synthesis of TE materials requires understanding the connection between various parameters involving synthesis conditions. In this direction, ML has great potential to control the synthesis conditions of novel TE materials. By learning existing synthesis information, ML can recommend efficient synthesis conditions with few trials. Thus, the ML can be used to understand complex relationships and predict optimal synthesis conditions with a high probability of success by employing existing initial synthesis data of TE material. Tang *et al.* have explored the feasibility of ML for guiding the $MoS_2$ synthesis by Chemical Vapor Deposition (CVD) and hydrothermal synthesis of carbon nanostructures for the first time[59]. Figure 7 shows the schematic of ML workflow for materials synthesis developed by Tang *et al.*[59]. This study has constructed an ML model for the general control of synthesis parameters for future experiments. The model could predict the probability of success and recommend the optimal synthesis conditions. Furthermore, a progressive adaptive model (PAM) to maximize the outcome of synthesis experiments through minimum trials has been introduced. This successful ML methodology for CVD of $MoS_2$ has the

**Table 1. The survey of this review refers to the descriptors, targets, and findings of the study related to ML in material chemistry**

| Publication | Sample source | Samples | Descriptors | Targets | Finding | Best ML algorithms |
|---|---|---|---|---|---|---|
| Juneja *et al.*, 2019[51] | Material project | 2,162 binary, ternary, and quaternary compounds | Maximum phonon frequency, integrated Gruneisen parameter up to 3 THz, average atomic mass, and volume of the unit cell | Analyzed the $\kappa_l$ for 120 materials in relation to the parameters | Low $\kappa_l$ materials between 0.13 and 0.98 ($CsK_2Sb$, TlI, $Ba_2BiAu$, SrTePd, $Ba_2SbAu$, TlBr, $Cs_2Se$, $PbI_2$, LiFeP, TlCl, $Ba_2AgSb$, $PbI_2$, LaCoTe) | GPR |
| Zhang *et al.*, 2018[55] | Previous reports | 93 binary semiconductors | Electronegativity, atomic radius, effective nuclear charge, Vander Waals radius, covalent radius, row number in the periodic table, block number, enthalpy of formation of gaseous atoms, ionization energy, and valence number | Proposed a method to increase prediction accuracy by incorporating property estimation in the feature space to establish ML models | Increased accuracy about 2.1% than before. Used to more accurately predict the $\kappa_l$ of materials having more than 0.1 $Wm^{-1}K^{-1}$ | Kernel ridge regression |
| Juneja *et al.*, 2020[52] | DFT + Boltztrap code | 2,838 compounds | Boiling point, melting point, specific heat, molar specific heat, molar volume, heat of fusion, heat of vaporization, Pauling electronegativity, first ionization energy, group and period in the periodic table, elemental thermal conductivity, atomic number, atomic mass, covalent radius, van der Waals radius, density, the average bond distance, average bond strength, volume per atom, volume of cell, and coordination number | Found the common parameters influencing electrical and thermal transport properties. Tested with 135 compounds | Common parameters driving chemical bonding such as electronegativity, bond strength, bond distance, volume, and coordination number | GPR |
| Juneja *et al.*, 2020[53] | DFT | 2,838 nonmetallic compounds after high throughput screening 185 compounds were analyzed | Boiling point, melting point, specific heat, molar specific heat, molar volume, heat of fusion, heat of vaporization, Pauling electronegativity, first ionization energy, group and period in the periodic table, elemental thermal conductivity, atomic number, atomic mass, covalent radius, van der Waals radius, density, the average bond distance, average bond strength, volume per atom, volume of cell, and coordination number | Found the correlation between the chemistry of bonding and $\kappa_l$ | The bond strength obtained was weak for high coordination numbers, and the bonding distance was large | GPR |
| Tewari *et al.*, 2020[50] | DFT | 315 compounds | lattice energy per atom, atom density, band gap, mass density, and oxygen ratio by transition metal atoms | Segregating and discovering material based on $\kappa_l$ | Identified low $\kappa_l$ among 315 transition metal oxides, values less than 5 $Wm^{-1}K^{-1}$ | XGBoost classifier and regression (cubist, kernel ridge, and Gaussian process) |

potential to be expanded to efficient synthesis of chalcogenide-based TE materials using CVD techniques.

Among chalcogenides, $Bi_2Te_3$-based materials show the best TE properties. Wang *et al.* have applied ML to optimize hot-extruded $Cu_xBi_2Te_{2.85+y}Se_{0.15}$ TE materials for the first time[60]. Principal component analysis (PCA) has been employed to characterize multiple variables. It has been found that the extrusion temperature and Cu content are the most important parameters in the $Cu_xBi_2Te_{2.85+y}Se_{0.15}$-based TE material processing design. This study developed artificial neural network (ANN), SVR, and RF regression models to analyze the relationships between processing conditions, microstructural information, and TE properties.

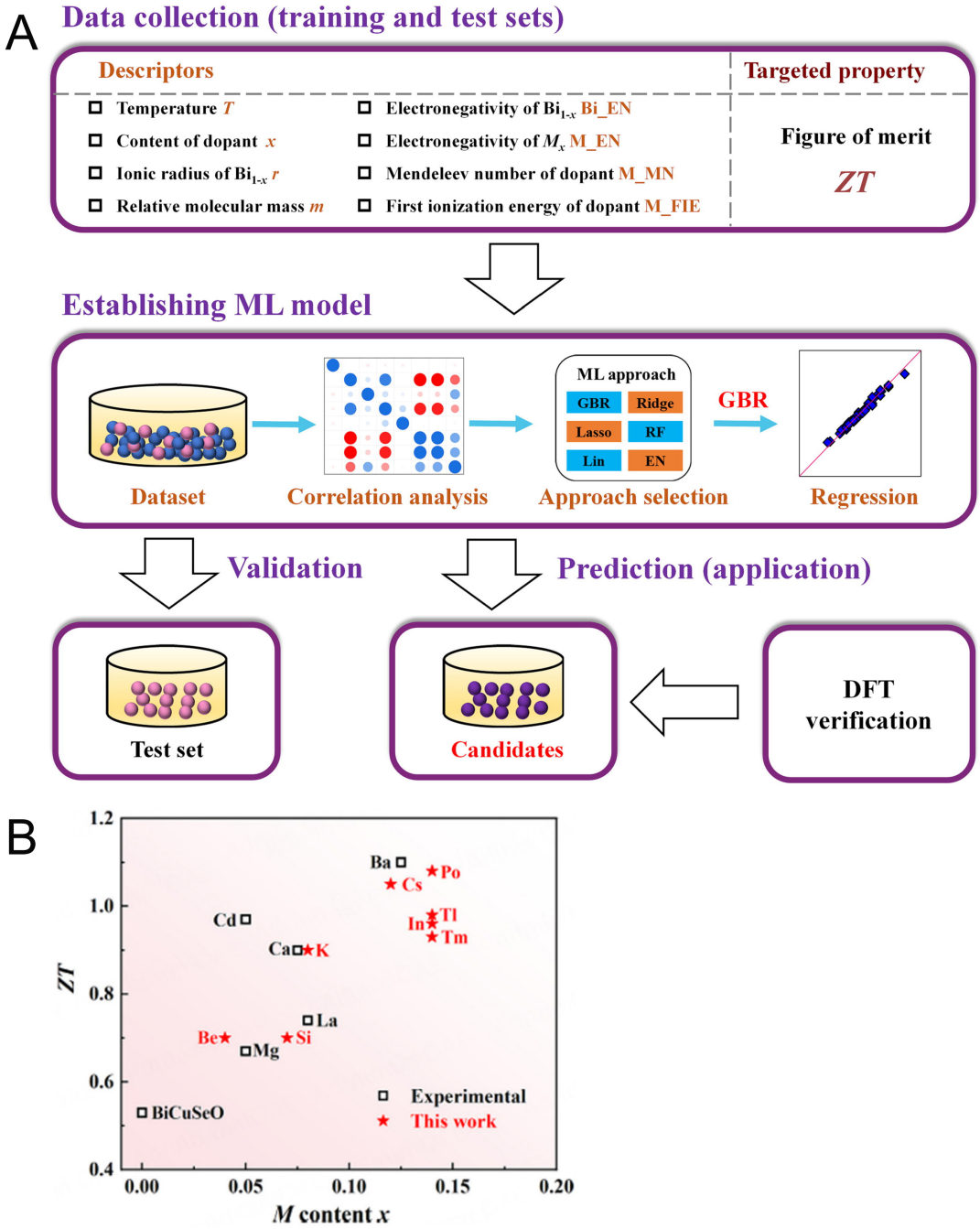


**Figure 6.** (A) Workflow and (B) identified ZT of doped elements through the ML approach[56]. (License Number 6015491243882).

The ML approach was used to understand the practical conditions involving coating, printing, and synthesizing the TE materials, as a result of which the time required to carry out the experiment and recover the material costs was lowered. BiSbTe is a well-known composition in TE applications exhibiting ZT greater than 1. However, predicting the suitable printing parameter through ML is an effective way to optimize the ZT. Song *et al.* employed a GPR-based ML model to predict the thermoelectric properties as a function of ink formulation and printing parameters in 3D printing[61]. Initially, the data was collected through experiment results to make ML algorithm. Considering the TE particle loadings and X-gum

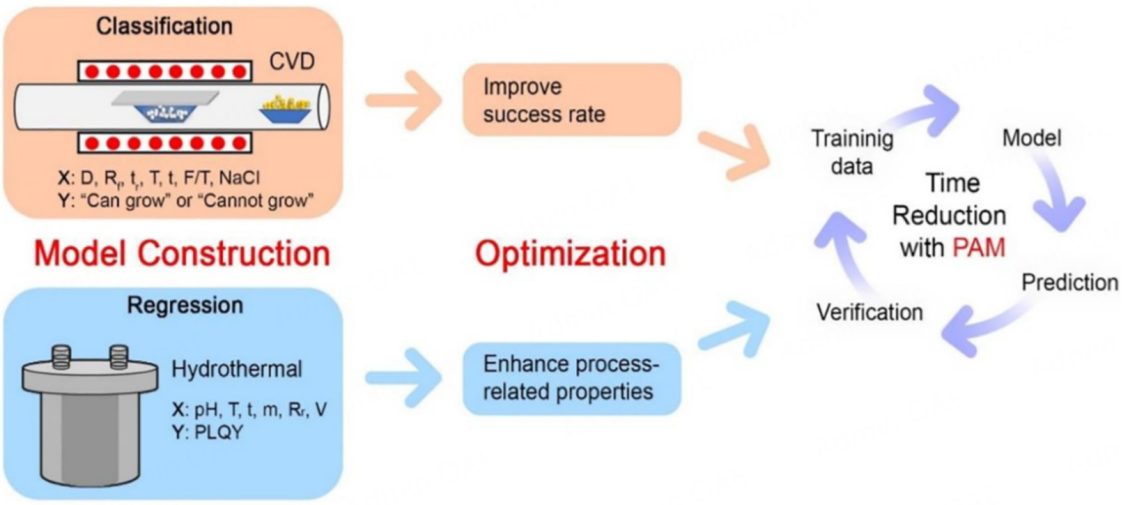


**Figure 7.** Schematics of ML-guided synthesis of materials by CVD and hydrothermal process. The three key steps of ML to material synthesis are model construction, optimization, and maximizing experimental outcomes. Adapted from Tang *et al.*[59]. (License Number 6015150101969).

concentrations for ink formulation, filament spacing, and stand-off distance for printing parameters, the ML approach optimizes the ZT of BiSbTe, as shown in Figure 8A. After round 4 optimization the predicted value through GPR was highly matched with the experimental value. Figure 8B shows the ML results. After applying the optimal condition, the material has a ZT of 1.3 at room temperature which was much higher than previous results.

$Bi_2Te_3$ is a well-known inorganic commercialized TE material operating at mid-temperature. The dopants Cu, Se, and Pb have been studied for $Bi_2Te_3$. Analyzing dopant concentration through ML is an effective way to get peak ZT. Alrebdi *et al.* developed support vector regression (SVR) (using both radial basis function and polynomial kernels) and DTR ML models to solve some real-world problems, including the effect of cu (metal), Se (non-metals), and Pb (toxic metal) on the values of $\kappa$[62]. The employed method results are almost close to the experimental data obtained from the literature; also, the minimum concentration of these Cu, Se, and Pb doping into $Bi_2Te_3$ exhibits the minimum thermal conductivity. In addition, the substrate temperature used for $Bi_2Te_3$ filmmaking shows considerable effect in the crystal formation during pulsed laser deposition (PLD), identified optimum substrate temperature through ML models matched with the reported experimental data. Identified through testing and training the datasets, the $R^2$ value was increased on trained data. The $R^2$ value of DTR was 1, indicating better fitting with the experimental value than SVR regression. This report gives the knowledge for future optimization processes using these ML models suitable to predict the optimum concentration of dopants and substrate temperature during PLD.

Sintering conditions influence the TE properties change while the optimum condition predicted through the ML approach reduces the number of iterations taken manually to increase the ZT. Also, more accurate knowledge about the sintering condition is not yet explored through experiments. Here, Headley *et al.* employed an ML approach to finding suitable sintering conditions for $Ag_{1.96}Se$ through an ultrafast flash sintering technique to achieve high performance[63]. The GPR model developed to identify optimum flash sintering variables includes voltage, pulse duration, pulse delay, number of pulses, and thickness related to power factor measured during the analysis based on Bayesian expected improvement. As the number of data increased, the GPR prediction uncertainty converged to the experimental measurement uncertainty. Finally, the optimum parameters were used to perform the experiment, and the maximum ZT was obtained

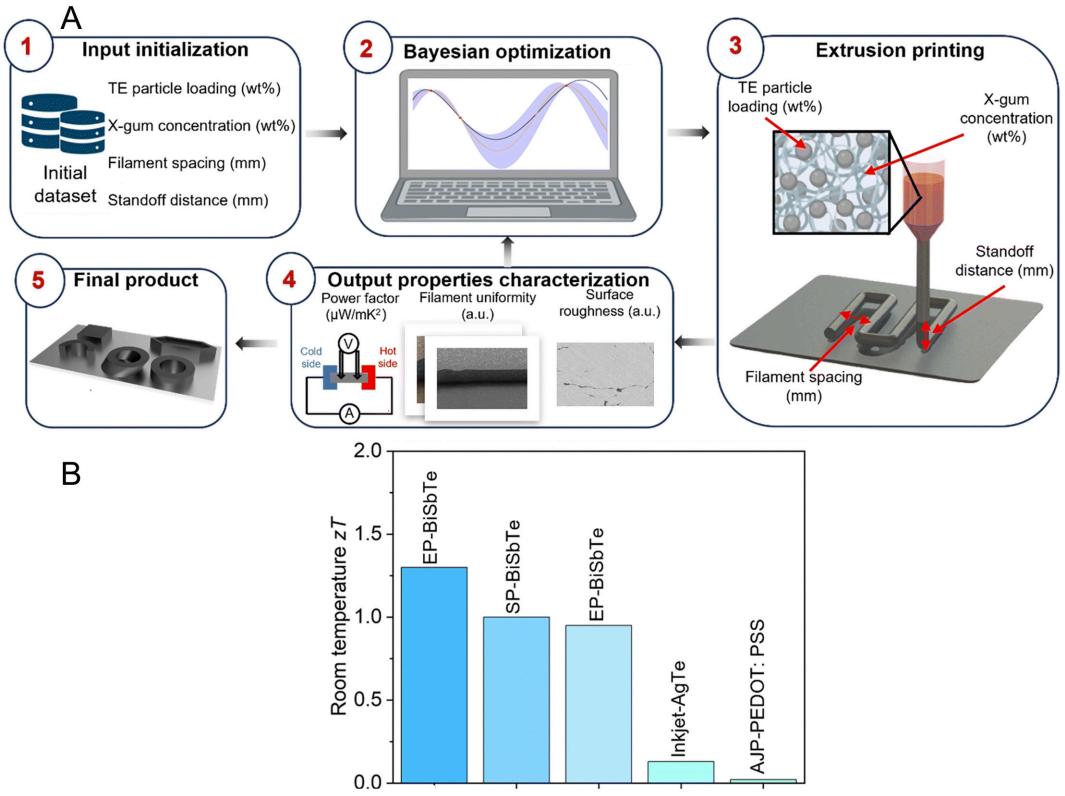


**Figure 8.** (A) Workflow of the ML-assisted extrusion printing of thermoelectric inks, including the four input variables listed in box 1 and three out properties of interests in box 4, (B) ZT of BiSbTe using the optimum condition identified through this ML model and reported values. Reproduced from Ref.[61] with permission from the Royal Society of Chemistry. (License under Creative Commons Attribution 3.0 Unported License).

within 1 s flash sintering time (optimized time). The sintering time was comparatively less than the previous sintering time.

Hou *et al.* applied ML to optimize the PF of $Al_2Fe_3Si_3$ by varying the composition of Al/Si[64]. Experimentally obtained data was used to train the ML model to predict unknown power factors. The commonly used anisotropic squared-exponential (SE) covariance function in GPR was chosen to describe the covariance between the feature variables of composition and temperature. Finally, the optimal ratio of 0.9 shows the increase of PF up to 40% at 510 K compared with the original composition. Headley *et al.* applied the ML approach to make a n-type $Bi_2Te_{2.7}Se_{0.3}$ under laser powder bed fusion (LPBF) processing[63]. The four steps are followed to predict the optimized LPBF-built complex geometries using an iterative augmented strategy, as shown in Figure 9. Initially, the new 13-line scan parameter (power and scan speed) combinations are predicted. Then, this parameter value was used as input for melt pool characterization. The width and depth of melt pool geometrical values were obtained, and these training datasets were used for ML. Again, 93-line scan parameter combinations were used to predict the melt pool geometry with uncertainties. Then, optimized parameter combinations were given to the LPBF-built to make $Bi_2Te_{2.7}Se_{0.3}$ with three geometries: the rectangular prism, hollow rectangle, and trapezoid. Integrating ML techniques helps to visualize and quickly understand changing melt pool dimensions concerning varied laser parameters.

### AI for TE materials synthesis

The first and only report during the writing of this article on the application of AI for TE materials synthesis

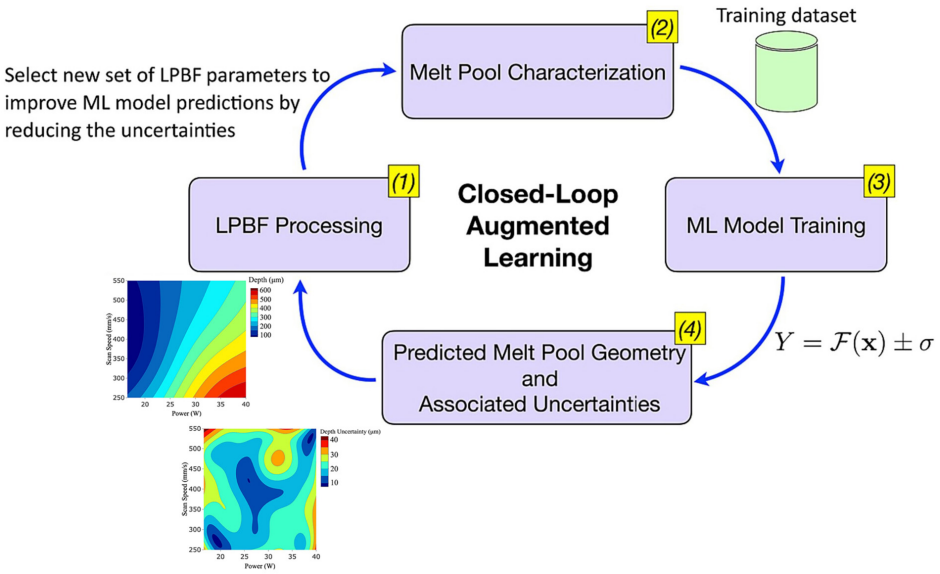


**Figure 9.** Showing the iterative augmented strategy to understand the melt pool geometrical (width and depth) using LPBF processing of n-type $Bi_2Te_{2.7}Se_{0.3}$ thermoelectric. A total of six iterations were performed with feedback[63]. (License number 6015220339088).

is by Na *et al.*[65]. This study converts chemical reaction formulas into a machine-readable format using a "synthesis graph". The synthesis graph describes the chemical formula in terms of the elements present in the starting and final materials, as shown in Figure 10. To predict the synthesis recipe of TE materials, a DNN-based architecture called the synthesis process encoder-decoder (SPENDE) has been developed [Figure 11]. Herein, based on the benchmark of synthesis dataset of 771 unique TE materials; first, a synthesis graph has been generated followed by calculation of graph embedding vector using a graph NN graph-based reaction encoder (GRE). Then, the NN operation sequence decoder (OSD) has achieved the prediction of each step's operation level. Finally, the preparation conditions are predicted by engineering conditions networks (ECNs) predicts. Thus, this architecture has successfully predicted the synthesis parameters involving grinding, heating, cooling, and sintering for synthesizing TE materials.

## CHALLENGES AND PERSPECTIVES

The rapid development of statistical methods has significantly influenced the discovery and design of TE material. This critically means that exciting new materials are being discovered; however, these materials are yet to be produced. ML and AI have the potential to learn patterns of synthesis design from a given data set of experimental synthesis procedures and then predict the outcome. Nevertheless, the design of synthesis parameters for TE materials is challenging because the sequence of reactions during the synthesis depends on many factors including choice of synthesis procedure and precursors. Therefore, the unavailability of an organized and comprehensive database of synthesis procedures of TE materials is a big challenge that needs to be overcome. For training the algorithm, converting available synthesis condition data and processing it into a suitable format for the algorithm is another obstacle. Finally, the synthesis parameters predicted by statistical optimization can be confirmed by experiments and further improved and developed.

Atomic properties are building blocks to construct the crystal system. Also, these parameters are related to the TE performance. Experimentally, finding this relation consumes more time and material cost. However, ML easily and quickly interprets this relation through the available data from materials projects or research articles. Interpretation of atomic properties with TE parameters tends to material classification which could help to identify new materials or compositions without any cost. Besides, the ML approach quickly

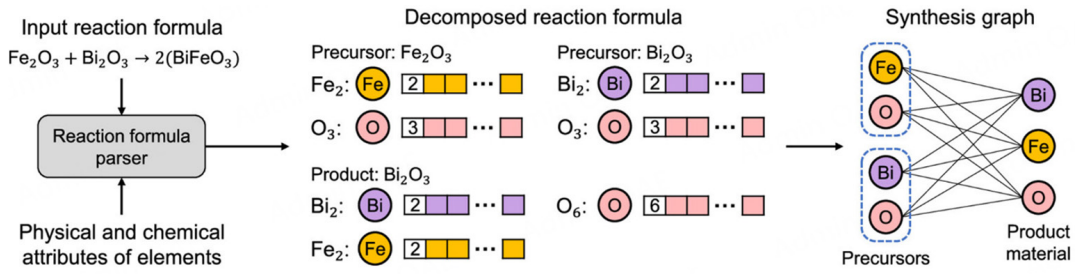


**Figure 10.** Example of a synthesis-graph for the conversion of a chemical reaction into a machine readable format developed by Na *et al.*[65]. (License CC-BY-NC-ND 4.0).

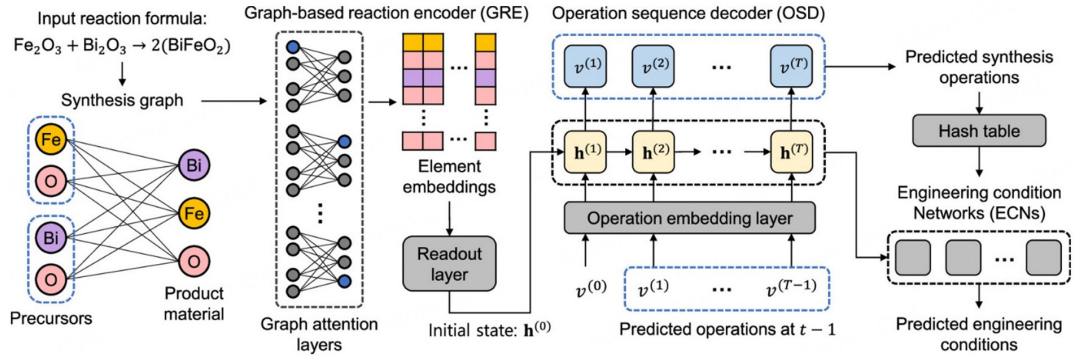


**Figure 11.** SPENDE architecture and its forward process for the prediction of synthesis sequence of a chemical reaction. Adapted from Na *et al.*[65] (License CC-BY-NC-ND 4.0).

identifies the exact atomic properties among many other atomic properties, helping to understand the material chemistry of the system. Further, it would help to develop the appropriate synthesis by controlling the identified atomic properties for high-performance TE material. Besides, controlling this parameter through experiments needs to be explored through the ML.

## CONCLUSIONS

In conclusion, statistical and data-driven methods such as DOE, ML, and AI have been reviewed to optimize and guide the synthesis of TE materials. Advanced statistical methods can simplify the complex TE materials synthesis process. In the traditional synthesis of TE materials, experiments are performed to measure the effects of experimental variables on responses. The optimization of the preparation of TE materials involves finding a combination of variables that gives the best results. Recent advances can guide the multi-variable synthesis of new TE material, improve the outcome of experiments, and save time. It has been demonstrated that the proposed methodologies may be extended to synthesize various categories of TE materials. Various atomic properties are involved in material chemistry. Also, finding atomic properties related to TE parameters extends to the material classification based on the performance. Materials classification based on the performance can be easily achieved through ML concerning atomic properties. Identifying suitable atomic properties helps develop a new composition or class of materials for TE without experimental cost.

## DECLARATIONS

### Acknowledgments

The authors gratefully acknowledge the financial support from the National Science and Technology

Council, Taiwan, R.O.C, under the contract NSTC 113-2221-E-006-195-MY3 for this study. This research is also supported in part by Higher Education Sprout Project, Ministry of Education to the Headquarters of University Advancement at National Cheng Kung University (NCKU).

### Authors' contributions

Conceptualization, formal analysis, funding acquisition, investigation, project administration, resources, software, supervision, writing - review & editing: Chen, W. H.
Data collection, formal analysis, investigation, writing - original draft: Aishwarya, K.; Sardar, K.

### Availability of data and materials

The data used for this study are confidential.

### Financial support and sponsorship

This work was financially supported by the National Science and Technology Council, Taiwan, R.O.C, under the contract NSTC 113-2221-E-006-195-MY3 for this study. This research is also supported in part by Higher Education Sprout Project, Ministry of Education to the Headquarters of University Advancement at National Cheng Kung University (NCKU).

### Conflicts of interest

All authors declared that there are no conflicts of interest.

### Ethical approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Copyright



## REFERENCES

1. Bell, L. E. Cooling, heating, generating power, and recovering waste heat with thermoelectric systems. *Science* **2008**, *321*, 1457-61. DOI PubMed
2. Beretta, D.; Neophytou, N.; Hodges, J. M.; et al. Thermoelectrics: from history, a window to the future. *Mater. Sci. Eng. R. Rep.* **2019**, *138*, 100501. DOI
3. DiSalvo, F. J. Thermoelectric cooling and power generation. *Science* **1999**, *285*, 703-6. DOI PubMed
4. Gorai, P.; Stevanović, V.; Toberer, E. S. Computationally guided discovery of thermoelectric materials. *Nat. Rev. Mater.* **2017**, *2*, 201753. DOI
5. Wang, T.; Zhang, C.; Snoussi, H.; Zhang, G. Machine learning approaches for thermoelectric materials research. *Adv. Funct. Mater.* **2020**, *30*, 1906041. DOI
6. Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Deng, Z.; Ong, S. P. A critical review of machine learning of energy materials. *Adv. Energy. Mater.* **2020**, *10*, 1903242. DOI
7. Recatala-Gomez, J.; Suwardi, A.; Nandhakumar, I.; Abutaha, A.; Hippalgaonkar, K. Toward accelerated thermoelectric materials and process discovery. *ACS. Appl. Energy. Mater.* **2020**, *3*, 2240-57. DOI
8. Wang, X.; Sheng, Y.; Ning, J.; et al. A critical review of machine learning techniques on thermoelectric materials. *J. Phys. Chem. Lett.* **2023**, *14*, 1808-22. DOI
9. Chen, W. H.; Carrera, U. M.; Kwon, E. E.; et al. A comprehensive review of thermoelectric generation optimization by statistical approach: Taguchi method, analysis of variance (ANOVA), and response surface methodology (RSM). *Renew. Sustain. Energy. Rev.* **2022**, *169*, 112917. DOI
10. Kucova, T.; Prauzek, M.; Konecny, J.; Andriukaitis, D.; Zilys, M.; Martinek, R. Thermoelectric energy harvesting for internet of things devices using machine learning: a review. *CAAI. Trans. on. Intell. Technol.* **2023**, *8*, 680-700. DOI
11. Song, K.; Tanvir, A. N. M.; Bappy, M. O.; Zhang, Y. New directions for thermoelectrics: a roadmap from high-throughput materials

discovery to advanced device manufacturing. *Small. Sci.* **2025**, *5*, 2300359. DOI

12. Deng, T.; Qiu, P.; Yin, T.; et al. High-throughput strategies in the discovery of thermoelectric materials. *Adv. Mater.* **2024**, *36*, e2311278. DOI
13. Chen, G.; Dresselhaus, M. S.; Dresselhaus, G.; Fleurial, J.; Caillat, T. Recent developments in thermoelectric materials. *Int. Mater. Rev.* **2003**, *48*, 45-66. DOI
14. Gayner, C.; Kar, K. K. Recent advances in thermoelectric materials. *Prog. Mater. Sci.* **2016**, *83*, 330-82. DOI
15. He, J.; Tritt, T. M. Advances in thermoelectric materials research: looking back and moving forward. *Science* **2017**, *357*, eaak9997. DOI PubMed
16. Yang, L.; Chen, Z. G.; Dargusch, M. S.; Zou, J. High performance thermoelectric materials: progress and their applications. *Adv. Energy. Mater.* **2018**, *8*, 1701797. DOI
17. Hasan, M. N.; Wahid, H.; Nayan, N.; Mohamed, A. M. S. Inorganic thermoelectric materials: a review. *Int. J. Energy. Res.* **2020**, *44*, 6170-222. DOI
18. Shi, X. L.; Zou, J.; Chen, Z. G. Advanced thermoelectric design: from materials and structures to devices. *Chem. Rev.* **2020**, *120*, 7399-515. DOI PubMed
19. Wei, J.; Yang, L.; Ma, Z.; et al. Review of current high-ZT thermoelectric materials. *J. Mater. Sci.* **2020**, *55*, 12642-704. DOI
20. Tan, G.; Zhao, L. D.; Kanatzidis, M. G. Rationally designing high-performance bulk thermoelectric materials. *Chem. Rev.* **2016**, *116*, 12123-49. DOI PubMed
21. Soleimani, Z.; Zoras, S.; Ceranic, B.; Shahzad, S.; Cui, Y. A review on recent developments of thermoelectric materials for room-temperature applications. *Sustain. Energy. Technol. Assess.* **2020**, *37*, 100604. DOI
22. Jarman, J.; Khalil, E. E.; Khalaf, E. Energy analyses of thermoelectric renewable energy sources. *Open. J. Energy. Effic.* **2013**, *2*, 143-53. DOI
23. Ryu, B.; Chung, J.; Kumagai, M.; et al. Best thermoelectric efficiency of ever-explored materials. *Iscience* **2023**, *26*, 106494. DOI PubMed PMC
24. Gaultois, M. W.; Sparks, T. D.; Borg, C. K. H.; Seshadri, R.; Bonificio, W. D.; Clarke, D. R. Data-driven review of thermoelectric materials: performance and resource considerations. *Chem. Mater.* **2013**, *25*, 2911-20. DOI
25. Wang, S.; Zuo, G.; Kim, J.; Sirringhaus, H. Progress of conjugated polymers as emerging thermoelectric materials. *Prog. Polym. Sci.* **2022**, *129*, 101548. DOI
26. Qin, Y.; Zhang, Q.; Chen, G. Organic borate doped carbon nanotube for enhancement of thermoelectric performance. *Carbon* **2021**, *182*, 742-8. DOI
27. Lindorf, M.; Mazzio, K. A.; Pflaum, J.; Nielsch, K.; Brütting, W.; Albrecht, M. Organic-based thermoelectrics. *J. Mater. Chem. A.* **2020**, *8*, 7495-507. DOI
28. Zhang, Y.; Wang, W.; Zhang, F.; et al. Soft organic thermoelectric materials: principles, current state of the art and applications. *Small* **2022**, *18*, e2104922. DOI
29. Deng, L.; Liu, Y.; Zhang, Y.; Wang, S.; Gao, P. Organic thermoelectric materials: niche harvester of thermal energy. *Adv. Funct. Mater.* **2023**, *33*, 2210770. DOI
30. Jin, H.; Li, J.; Iocozzia, J.; et al. Hybrid organic-inorganic thermoelectric materials and devices. *Angew. Chem. Int. Ed.* **2019**, *58*, 15206-26. DOI
31. Selvaratnam, B.; Koodali, R. T. Machine learning in experimental materials chemistry. *Catal. Today.* **2021**, *371*, 77-84. DOI
32. Williamson, E. M.; Brutchey, R. L. Using data-driven learning to predict and control the outcomes of inorganic materials synthesis. *Inorg. Chem.* **2023**, *62*, 16251-62. DOI PubMed PMC
33. Baum, F.; Pretto, T.; Köche, A.; Santos, M. J. L. Machine learning tools to predict hot injection syntheses outcomes for II-VI and IV-VI quantum dots. *J. Phys. Chem. C.* **2020**, *124*, 24298-305. DOI
34. Kyratsi, T.; Ioannou, M. Thermoelectric properties of hot-pressed β-$K_2Bi_8Se_{13-x}S_x$ materials. *J. Electron. Mater.* **2013**, *42*, 1604-11. DOI
35. Kanatzia, A.; Papageorgiou, C.; Lioutas, C.; Kyratsi, T. Design of ball-milling experiments on $Bi_2Te_3$ thermoelectric material. *J. Electron. Mater.* **2013**, *42*, 1652-60. DOI
36. Nuthongkum, P.; Sakulkalavek, A.; Sakdanuphab, R. RSM base study of the effect of argon gas flow rate and annealing temperature on the [Bi]:[Te] ratio and thermoelectric properties of flexible Bi-Te thin film. *J. Electron. Mater.* **2017**, *46*, 2900-7. DOI
37. Khumtong, T.; Sakulkalavek, A.; Sakdanuphab, R. Empirical modelling and optimization of pre-heat temperature and Ar flow rate using response surface methodology for stoichiometric $Sb_2Te_3$ thin films prepared by RF magnetron sputtering. *J. Alloys. Compd.* **2017**, *715*, 65-72. DOI
38. Zhang, Y.; Zhang, Q.; Chen, G. Carbon and carbon composites for thermoelectric applications. *Carbon. Energy.* **2020**, *2*, 408-36. DOI
39. Jagadish, P. R.; Khalid, M.; Amin, N.; Li, L. P.; Chan, A. Process optimisation for n-type $Bi_2Te_3$ films electrodeposited on flexible recycled carbon fibre using response surface methodology. *J. Mater. Sci.* **2017**, *52*, 11467-81. DOI
40. Sam, S.; Sreypich, S.; Abad, A.; Gan Lim, L.; Santos, G. N. Fabrication and characterization of PbSnTe crystals for thermoelectric applications. *J. Comput. Innov. Eng. Appl.* **2022**, *2*, 1-2. Available from: https://www.researchgate.net/publication/360112987_Fabrication_and_Characterization_of_PbSnTe_Crystals_for_Thermoelectric_Applications?enrichId=rgreq-db9922f1349b27a902702d24605a71362-XXX&enrichSource=Y292ZXJQYWdlOzM2MDExMjk4NztBUzoxMTQ3Njg0MTMxOTk5NzQ5QDE2NTA2NDA2MjgzNzE%3D&el=1_x_2&_esc=

publicationCoverPdf [Last accessed on 4 Jun 2025]

41. Lange, R. G.; Carroll, W. P. Review of recent advances of radioisotope power systems. *Energy. Convers. Manag.* **2008**, *49*, 393-401. DOI
42. Basu, R.; Bhattacharya, S.; Bhatt, R.; et al. Improved thermoelectric performance of hot pressed nanostructured n-type SiGe bulk alloys. *J. Mater. Chem. A.* **2014**, *2*, 6922. DOI
43. Ahmad, S.; Singh, A.; Basu, R.; et al. Optimization of thermoelectric properties of mechanically alloyed p-type SiGe by mathematical modelling. *J. Electron. Mater.* **2019**, *48*, 649-55. DOI
44. Karthikeyan, V.; Surjadi, J. U.; Li, X.; et al. Three dimensional architected thermoelectric devices with high toughness and power conversion efficiency. *Nat. Commun.* **2023**, *14*, 2069. DOI PubMed PMC
45. Maduabuchi, C. Thermo-mechanical optimization of thermoelectric generators using deep learning artificial intelligence algorithms fed with verified finite element simulation data. *Appl. Energy.* **2022**, *315*, 118943. DOI
46. Maduabuchi, C.; Eneh, C.; Alrobaian, A. A.; Alkhedher, M. Deep neural networks for quick and precise geometry optimization of segmented thermoelectric generators. *Energy* **2023**, *263*, 125889. DOI
47. Ameenuddin Irfan, S.; Irshad, K.; Algahtani, A.; et al. Machine learning-based modeling of thermoelectric materials and air-cooling system developed for a humid environment. *Mater. Express.* **2021**, *11*, 153-65. DOI
48. Gulevich, D.; Nabiev, I.; Samokhvalov, P. Machine learning-assisted colloidal synthesis: a review. *Mater. Today. Chem.* **2024**, *35*, 101837. DOI
49. Iwasaki, Y.; Takeuchi, I.; Stanev, V.; et al. Machine-learning guided discovery of a new thermoelectric material. *Sci. Rep.* **2019**, *9*, 2751. DOI PubMed PMC
50. Tewari, A.; Dixit, S.; Sahni, N.; Bordas, S. P. A. Machine learning approaches to identify and design low thermal conductivity oxides for thermoelectric applications. *Data-Centric. Eng.* **2020**, *1*, e8. DOI
51. Juneja, R.; Yumnam, G.; Satsangi, S.; Singh, A. K. Coupling the high-throughput property map to machine learning for predicting lattice thermal conductivity. *Chem. Mater.* **2019**, *31*, 5145-51. DOI
52. Juneja, R.; Singh, A. K. Unraveling the role of bonding chemistry in connecting electronic and thermal transport by machine learning. *J. Mater. Chem. A.* **2020**, *8*, 8716-21. DOI
53. Juneja, R.; Singh, A. K. Guided patchwork kriging to develop highly transferable thermal conductivity prediction models. *J. Phys. Mater.* **2020**, *3*, 2. DOI
54. Li, W.; Liu, M. Interpretable machine learning workflow for evaluating and analyzing the performance of high-entropy GeTe-based thermoelectric materials. *ACS. Appl. Electron. Mater.* **2023**, *5*, 4523-33. DOI
55. Zhang, Y.; Ling, C. A strategy to apply machine learning to small datasets in materials science. *NPJ. Comput. Mater.* **2018**, *4*, 81. DOI
56. He, Z.; Peng, J.; Lei, C.; Xie, S.; Zou, D.; Liu, Y. Prediction of superior thermoelectric performance in unexplored doped-BiCuSeO via machine learning. *Mater. Des.* **2023**, *229*, 111868. DOI
57. Minhas, H.; Jena, M. K.; Sharma, R. K.; Pathak, B. Machine learning-driven inverse design and role of dopant for tuning thermoelectric efficiency. *ACS. Appl. Electron. Mater.* **2024**, *6*, 5815-26. DOI
58. Parse, N.; Pongkitivanichkul, C.; Pinitsoontorn, S. Machine learning approach for maximizing thermoelectric properties of BiCuSeO and discovering new doping element. *Energies* **2022**, *15*, 779. DOI
59. Tang, B.; Lu, Y.; Zhou, J.; et al. Machine learning-guided synthesis of advanced inorganic materials. *Mater. Today.* **2020**, *41*, 72-80. DOI
60. Wang, Z.; Adachi, Y.; Chen, Z. C. Processing optimization and property predictions of hot-extruded Bi-Te-Se thermoelectric materials via machine learning. *Adv. Theory. Simul.* **2020**, *3*, 1900197. DOI
61. Song, K.; Xu, G.; Tanvir, A. N. M.; et al. Machine learning-assisted 3D printing of thermoelectric materials of ultrahigh performances at room temperature. *J. Mater. Chem. A.* **2024**, *12*, 21243-51. DOI
62. Alrebdi, T.; Wudil, Y.; Ahmad, U.; Yakasai, F.; Mohammed, J.; Kallas, F. Predicting the thermal conductivity of $Bi_2Te_3$-based thermoelectric energy materials: a machine learning approach. *Int. J. Therm. Sci.* **2022**, *181*, 107784. DOI
63. Headley, C. V.; Herrera del Valle, R. J.; Ma, J.; et al. The development of an augmented machine learning approach for the additive manufacturing of thermoelectric materials. *J. Manuf. Processes.* **2024**, *116*, 165-75. DOI
64. Hou, Z.; Takagiwa, Y.; Shinohara, Y.; Xu, Y.; Tsuda, K. Machine-learning-assisted development and theoretical consideration for the $Al_2Fe_3Si_3$ thermoelectric material. *ACS. Appl. Mater. Interfaces.* **2019**, *11*, 11545-54. DOI PubMed
65. Na, G. S. Artificial intelligence for learning material synthesis processes of thermoelectric materials. *Chem. Mater.* **2023**, *35*, 8272-80. DOI