ARTICLE

High thermoelectric power factor in Ni-Fe alloy for active cooling applications

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Shuai Li,[a] Sree Sourav Das,[b]† Haobo Wang,[a]† Kacper Pryga,[c] Sujit Bati,[d] Bartlomiej Wiendlocha,[c] Junichiro Shiomi,[e] Jerrold A. Floro,[a] Prasanna V. Balachandran,[a] and Mona Zebarjadi [a,b]

Metallic thermoelectric materials are promising candidates for active cooling applications, where high thermal conductivity and a high thermoelectric power factor are essential to maximize effective thermal conductivity. While metals inherently possess high thermal and electrical conductivities, they typically exhibit low Seebeck coefficients. In this work, we create a database of the Seebeck coefficient of binary metallic alloys and apply machine-learning techniques to identify alloys with large Seebeck coefficients. Specifically, we identify Ni-Fe as a promising candidate for active cooling around room temperature. We then fabricate Ni-Fe ingots and demonstrate thermoelectric power factor values as high as 120 μW/cm·K² at 200 K for these stable alloys, which are composed of cost-effective and abundant elements. Furthermore, we show that the effective thermal conductivity of these alloys, under small temperature differences, can exceed that of pure copper at 250-400 K temperature range.

1. Introduction

With the increasing density of transistors and operating frequencies in integrated circuits (ICs), driven by rapid advancements in semiconductor technologies, efficient heat dissipation has become an increasingly critical challenge. Inadequate heat management impairs the performance of these densely packed circuits and jeopardizes their reliability. Conventional cooling techniques, such as passive heat sinks and fluid-based cooling systems, often struggle to meet the efficiency, size, and design requirements of ICs. In response, novel approaches—such as active cooling based on the thermoelectric (TE) effect—have emerged as promising solutions for thermal management.

TE materials, known for their ability to convert thermal gradients into electrical energy and vice versa, offer versatile applications through both the Seebeck and Peltier effects. These materials have long been studied for power generation and refrigeration, with the performance of TE devices governed by the dimensionless figure of merit, $zT = \frac{\sigma\alpha^2}{\kappa}T$, where σ is the electrical conductivity, α is the Seebeck coefficient, T is the temperature, and κ is thermal conductivity. We note that κ is the passive thermal conductivity in the absence of electric current. Improvement of zT requires strategies to increase the TE power factor ($PF = \sigma\alpha^2$) and/or to decrease the κ.[1–5] However, recent advances have expanded the role of TE materials to include active cooling modes[6–9], where the Peltier current can actively enhance the passive heat transfer. Under optimum current conditions, and when a TE module with a length L is placed between a hot object characterized by T_H and a cold heat sink at T_C, Peltier cooling (flux) can be expressed as $J_{peltier} = \frac{\sigma\alpha^2 T_H^2}{2L}$. One can therefore combine passive and active heat flux, where this unique mode of operation is characterized by the concept of effective thermal conductivity $\kappa_{eff} = \left(\kappa + \frac{\sigma\alpha^2 T_H^2}{2\Delta T}\right)$, combines the passive ($\kappa$) and active (Peltier) components to maximize the heat flux, opening new opportunities for TE application in thermal management. In our prior work, we have demonstrated cooling devices working based on a large κ_{eff}.[6]

Traditional TE materials with low κ are not suitable for active cooling. Instead, metallic TE materials are promising due to their inherently high electrical and thermal conductivities, originating from the high concentration of free electrons, n. The main disadvantage of metals is their generally lower Seebeck coefficient compared to semiconductors. This trend can be explained by the Mott Formula.

$$S = -\frac{\pi^2 k_B^2 T}{3q\sigma(\mu)}\frac{d\sigma(E)}{dE}\Big|_{E=\mu} \quad (1)$$

Where k_B is the Boltzmann constant, T is the temperature, q is the charge of the electron, E is energy, μ is the chemical potential. σ is differential conductivity or transport function

[a] Department of Materials Science and Engineering, University of Virginia, Charlottesville, VA, 22904, United States
[b] Electrical and Computer Engineering Department, University of Virginia, Charlottesville, VA, 22904, United States
[c] Faculty of Physics and Applied Computer Science, AGH University of Krakow, Aleja Mickiewicza 30, 30-059 Krakow, Poland
[d] Department of Physics, University of Virginia, Charlottesville, VA, 22904, United States
[e] Department Of Mechanical Engineering, University of Tokyo, Tokyo, Japan
† Sree Sourav Das and Haobo Wang contributed equally to this paper.
Supplementary Information available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

and can be expressed as $\sigma(E) = \frac{q^2}{3} DOS(E) v^2(E) \tau(E)$, DOS is the density of states, τ represents the relaxation time, and v is the group velocity. A high *DOS* at the Fermi level leads to an increased carrier density, which in turn enhances both electrical and thermal conductivity. However, according to the Mott formula (Eq. 1), a large *DOS* is associated with a reduced Seebeck coefficient, hence the lower Seebeck coefficient of the metals.

In this work, we lay out a search for binary metallic alloys for active cooling applications. Since metallic alloys inherently have a large electrical and thermal conductivity, we focus on finding alloys with large Seebeck coefficients. Historically, materials optimization has relied on traditional trial-and-error methods, guided by physical and chemical insights. However, these approaches are often time-consuming, inefficient, and costly.[10] To accelerate the optimization process, we employ a machine learning (ML) approach. By utilizing a data-driven model, our method efficiently explores a broad range of metallic alloys, focusing on different atomic concentrations, temperatures, and predicted Seebeck coefficients. In what follows, we present the material database that we built based on binary solid-solution metallic alloys. Using this database, we selected and optimized three alloys and finally, validated the results experimentally for the Ni-Fe alloy system.

2. Results and Discussion

2.1. Material Selection using Machine Learning

Near room temperature, a large list of pure metals, including Ag, Al, Au, Cd, Cs, Cu, Dy, In, Ir, Mg, Nb, Pb, Rh, Sn, Sr, and Ru, show absolute Seebeck coefficient values below 5 $\mu V/K$.[11] A few elemental metals, e.g., Co[6,12], Fe[13], and Ni[14] have larger Seebeck coefficient values (|S| ~ 20$\mu V/K$).[11] In the case of Ni, the large Seebeck coefficient can be attributed to the sharp slope of the *DOS* at the Fermi energy due to the partially filled d-orbital[15], which leads to a high Seebeck (Eq. 1). Due to the inherent magnetization of these elements (Ni, Co & Fe), part of their Seebeck coefficient has been attributed to the magnon-drag effect wherein the magnon heat flux drags along the electronic charge carriers. Watzman et al. [16] discussed two magnon-drag contributions to the Seebeck coefficient: the hydrodynamic contribution and the spin-motive force contribution. They demonstrated that magnon-drag is the dominant component of the Seebeck coefficient of iron and cobalt.

Binary alloys of transition metals are shown to be good candidates to further increase the TE power factor of metallic systems. Examples include Cu-Ni[8,17], Au-Ni[15], Fe-Ni[18,19], Cr-Mn[20], Pd-Ag[21], and Cr-Fe[22]. We formed a database of experimentally reported Seebeck coefficients of binary metallic alloys, the majority of which are derived from the Landolt-Bornstein database.[23] A part of this database for temperatures up to 400 K is shown in Figure 1. In several cases, the Seebeck coefficient and the power factor of the alloy are larger than both parent elements, which can be attributed to either changes in the density of states or modifications of the scattering rates. Here we highlight three of the studied solid-solution alloys with large Seebeck coefficient values, Pd-Ag, Cu-Ni, and Au-Ni. In all three cases, the alloy demonstrates a higher Seebeck coefficient compared to both the host and solute. The Pd-Ag alloy exhibits a peak Seebeck coefficient of approximately 40 μV/K at 300 K, reaching ~ 80 μV/K at 1300 K with 55% Pd.[21] These values are larger than pure Pd, which has a negative Seebeck coefficient in the 300K to 1300K range, and pure Ag with positive values below 10μV/K. However, the limited availability of Pd and the cost of the elements restricts its widespread use. Constantan (Cu-Ni) alloy is composed of abundant and low-cost elements and is easy to synthesize.[24] Constantan is reported to have a TE power factor of 40 $\mu \mathrm{W\,cm^{-1}\,K^{-2}}$ at 300 K and 102 $\mu \mathrm{W\,cm^{-1}\,K^{-2}}$ at 873 K.[17] Constantan is also studied in the context of active cooling and using additive manufacturing for industrial applications.[8] Au-Ni, in contrast, is expensive and metastable. However, due to its very large TE power factor worthy of investigation. Garmroundi et al.[15] reported a Seebeck coefficient of 94$\mu V/K$ for a quenched, metastable single face-centered cubic (FCC) Ni-Au alloy at 1000 K, resulting in an ultra-high peak power factor of 340 $\mu \mathrm{W\,cm^{-1}\,K^{-2}}$ in $Ni_{0.1}Au_{0.9}$ sample at 560K. This large power factor is hypothesized to arise from the selective scattering of s-electrons into localized d-states, which induces strong energy-dependent scattering rates $\tau(E)$ and enhances the slope of $\sigma(E)$ near the Fermi-level, thereby increasing the Seebeck coefficient (see Eq. 1).

Ni-Fe alloys have also been studied due to their significance in geology, meteoritics, and material science.[25–33] There are old and scattered studies reporting the Seebeck coefficient data of Ni-Fe alloys as a function of composition and temperature [18,19,34], showing the peak Seebeck of Ni-Fe alloys can reach -50 μV/K with 53.8 at% at 300 K [19] and -46 μV/K with 40 at%[18] of Ni concentration at 200 K. Figure 2(c-d) highlights that while the Fe-Ni system has been extensively studied between 300 and 1200 K, only a limited number of compositions have been investigated below 300 K. Similarly, Cr-Mn and Cr-Fe alloys were also examined but within a specific compositional range, as depicted in Figures S2(a) and S2(b), respectively. This research gap drives our focus on these three binary systems, aiming to enhance their performance through further composition optimization from low temperatures to room temperature ranges and explore the potential alloys for active cooling applications.

To predict the Seebeck coefficient, we used the database with selected features (see methods) to train several ML models. We employed three distinct types of ML models: a linear model based on Least Absolute Shrinkage and Selection Operator (LASSO)[35]regression, tree-based models (Extreme Gradient Boosting (XGBoost)[36] and Random Forest (RF)[37], and a kernel-based model (Support Vector Regression (SVR)[38]. Based on the accuracy-interpretability trade-off[39], linear models offer higher interpretability, while kernel-based models provide greater accuracy at the cost of interpretability. Tree-based models fall between these two extremes, balancing accuracy and interpretability.

Figure 3 presents a comparative analysis of the predicted vs. actual Seebeck coefficients across different models. The low R^2

and high MSE values in both the training and test sets for LASSO indicate that the linear model failed to capture the complex relationships between the input features and Seebeck coefficients. In contrast, XGB, RF, and SVR demonstrated a strong predictive performance, achieving an R^2 of 0.99 on both training and independent test sets, highlighting their ability to model the non-linear dependencies. After successfully evaluating the prediction performance of the models on the test data, we applied the optimized models (XGB, RF, and SVR) to predict the Seebeck coefficients of many binary alloys and identified Ni-Fe, Cr-Mn, and Cr-Fe alloys as our focus across the entire compositional range and from 50 to 305 K. The final Seebeck coefficient values were determined by averaging the predictions from all three models, with detailed predictions provided in Table S3. Based on these averaged results, the Ni-Fe system exhibited the highest Seebeck coefficient of 42.3 μV/K at 59 at% Ni and 305 K. Similarly, the Cr-Mn system reached a peak Seebeck coefficient of 34.6 μV/K at 10 at% Mn and 305 K, while the Cr-Fe system achieved 36.4 μV/K at 2 at% Fe and 155 K. Heat maps illustrating the Seebeck coefficients of the Cr-Mn and Cr-Fe systems are shown in Figure S5. Among these three systems, Ni-Fe presents the highest Seebeck coefficient. In addition, experimental validation of Cr-based systems presents challenges due to chromium's high reactivity with steel milling jars, which can alter the sample composition and degrade the performance.[40,51] Therefore, in this study, we focus on the Ni-Fe binary alloy system, given its high Seebeck coefficient near room temperature as predicted by the ML models. Additionally, this system is notable for its low cost, scalability for industrial applications, ease of synthesis, and exceptional mechanical durability.[41,42] These characteristics make Ni-Fe alloys highly suitable for practical TE device applications. According to the heat map of the predicted Seebeck coefficients (Figure 4), while the peak composition is $Ni_{56}Fe_{44}$, the alloy retains a Seebeck coefficient above -40 μV/K within the 48–62 atomic % Ni range around from 250 K to room temperature. This indicates a compositional window where the material could exhibit high TE performance. By identifying the peak Seebeck coefficient and compositional range, the ML model provides a focused direction for experimental exploration, minimizing the need for testing across all possible compositions.

2. 2. Experimental Validation

Next, we present our experimental approach to the power factor and effective thermal conductivity characterization of the Ni-Fe alloys in the 50-400K temperature range inspired by the ML predictions. Further, we present the microstructure of the as-arc-melted sample and the homogeneity of the solid-solution alloy at microscales for active cooling applications.

Ni_xFe_{1-x} samples with x atomic percentage ranging from 30-70 were prepared using arc-melting, see methods. According to the phase diagram, Ni-Fe alloys form an FCC Ni-Fe solid solution within the probed composition range at elevated temperatures. Due to slow diffusion [28–31,43], decomposition of γ to α + $FeNi_3$ is unlikely under our experimental conditions, resulting in γ phase solid-solution.

The X-ray diffraction (XRD) data of all Ni-Fe alloy samples are shown in Figure 5. The bulk XRD measurements are performed on the vertical cross-section of the as-arc-melted ingots. Going from bottom to top, Ni concentration in the samples increases. All samples exhibit a consistent series of diffraction peaks, corresponding to the (111), (200), (220), and (311) crystallographic planes of the FCC structure. The variation of intensities of peaks among samples can be attributed to non-random distributions of grain orientations at the section surface, which is confirmed by the SEM results. The orange and black lines at the bottom indicate the reference peak positions for pure Ni and Fe in the FCC structure. The peaks of the alloys are positioned between the reference peaks for pure Ni and Fe, indicative of solution formation. As the Ni concentration increases, the peaks shift to higher 2θ values, indicating smaller lattice parameters.[44] The XRD results confirmed that the prepared Ni-Fe samples are single-phase polycrystalline.

Figure 6 presents the backscattered electron image of the vertical cross-section of the as-arc-melted $Ni_{55}Fe_{45}$ sample. The different greyscale regions represent the varying crystal orientations of the grains, demonstrating the polycrystalline nature of the sample. At the top of the sample, a prominent needle-shaped bubble is observed, likely formed from the degassing of the powders during the melting process. Several smaller bubbles can also be seen in the upper section of the cross-section. To avoid these bubbles, subsequent transport measurements were performed on samples cut from the center of the lower portion, where grains are more uniform.

The SEM image clearly illustrates the distribution of grain size and shape across the sample. The bottom of the sample, which was in contact with the water-cooled copper plate of the arc-melter, experienced a higher cooling rate, resulting in smaller grains (~ 100 μm). In contrast, the upper section contains large, elongated grains measuring up to several millimeters in length. Further EDS mapping of the highlighted area was conducted to characterize the composition of different grains. The black dots visible in the enlarged image are colloidal silica residues from the sample polishing process. The atomic composition data in Figure 6 b's summary table confirms the homogeneous composition across the grains, consistent with the stoichiometric ratio of the starting powders. As shown in Figure 6c, the line scan reveals that the composition remains uniform both within and across grains. Additional SEM/EDS characterizations (Supplementary) performed on different samples and in different areas and orientations support that the samples are homogeneous and consistent with the measured composition. Due to anisotropic alignment of the grains, for the TE measurements, we only used the central-bottom part of the arc-melted sample, which is visually isotropic. However, due to the large grain sizes, we expect a minimal grain boundary effect on transport properties.

Figure 7 summarizes the TE measurements performed on the arc-melted Ni-Fe alloy samples.

Alloys with a composition range of 45 to 70 atomic % Ni have an absolute value of the Seebeck coefficient which is up to 2.5 times greater than that of pure Ni or Fe.[11] The peak Seebeck coefficient varies with composition, with the highest values observed in $Ni_{55}Fe_{45}$ and $Ni_{45}Fe_{55}$, both reaching -52 $\mu V/K$. This is consistent with the ML prediction presented earlier, where 45-55% Ni was identified as the composition with the highest Seebeck value. The Seebeck coefficient's dependence on composition changes with temperature: at lower temperatures (<200K), the absolute value of the Seebeck coefficient decreases with increasing Ni content. However, this trend does not hold at intermediate temperatures (200K to 400K).

A previous work[19] observed a similar concentration dependence of the Seebeck coefficient. They attributed the trend at the higher temperatures to the concentration fluctuation within their samples, which is not supported by the SEM/EDS results in this paper.

While ML predictions are consistent with experimental data, they cannot explain the origin of the large observed Seebeck coefficient. To understand this, we have computed the band structure using the first-principles methods. The details of DOS calculations for one unit cell of $Ni_{50}Fe_{50}$ are shown in the supplementary materials, wherein we have shown that the slope of the DOS at the Fermi level (application of the Mott formula) does not correctly predict the sign of the Seebeck coefficient. We have further expanded our calculations to the full-potential Korringa–Kohn–Rostoker method combined with the coherent potential approximation (KKR-CPA)[45,46], ensuring the correct description of alloy band structure, magnetism, and the disorder-induced scattering in the system. Densities of states and Bloch spectral density functions, which describe the electronic dispersion relations smeared due to electron scattering, are shown and discussed in the supplementary materials. Further on, the transport properties were determined by computing the energy-dependent conductivity function $\sigma(E)$ from the Kubo-Greenwood formalism.[47–50] As experimentally determined[51] this alloy is ferromagnetic with a high Curie temperature of 789 K, thus, calculations were done in a ferromagnetic state. Based on the computed $\sigma(E)$ function, the thermopower was calculated (see supplementary materials for further details). What is important to underline here is that this method has successfully predicted the Seebeck coefficient of other metallic alloys, including Ni-Au[52], Ni-Cu[53,54], and Pd-Ag[55], where resonant scattering effects are important. However, as shown in Figure 7c, in the case of $Ni_{50}Fe_{50}$ alloy, while the KKR-CPA method predicts the Seebeck coefficient sign correctly for magnetic calculations, the absolute value is much smaller than the experimentally measured values. Hence, the sole electronic structure and electron scattering on the atomic potentials do not explain the large thermopower values, and other energy-dependent scattering rates are needed along with magnon-drag contributions to fully understand the Seebeck values of this alloy system. The fact that additional scattering mechanisms (beyond electron-phonon) are present in this system is also confirmed by the difference in calculated and experimental residual resistivities. The calculated value at zero Kelvin is equal to about 3.3 μΩ·cm, whereas the experimental value extrapolated to zero Kelvin is larger, being about 11.9 μΩ·cm for $Ni_{50}Fe_{50}$.

The resistivity increases with higher Fe content. The resistivities of the Ni-Fe alloys range from 5.60 μΩ·cm to 70 μΩ·cm, highlighting the highly metallic nature of these alloys. Combining high Seebeck coefficients for these alloys with their low resistivity, the peak power factor reaches 120 μW/cm·K² for both $Ni_{60}Fe_{40}$ and $Ni_{55}Fe_{45}$. This is larger than both Ni and Fe parent metals[12] and is 20% higher than the peak values reported at 750K in previous studies on Cu-Ni alloys.[8,17] In Figure 7, we also compare the power factor to other binary metals with large power factors including PdAg and PdAu alloys.[21,56] In temperature ranges slightly below room temperature (i.e. 200K to 300K), there are not many candidates with extremely large TE power factors (i.e., above 100 μW/cm·K²). Commonly used TE materials in this temperature range include bismuth-tellurium-antimony-selenium-based materials, which generally, have power factor values well below 100 μW/cm·K², with a recent work highlighting a record high value of 63 μW/cm·K² in this class of materials.[57] Au-Ni is reported to have a power factor slightly below 300 μW/cm·K².[15] However, its cost and instability are not favorable. Single-crystal $YbAl_3$ has a power factor slightly below 200 μW/cm·K² at room temperature.[58,59] Other examples include low-dimensional materials such as nm thin FeSe[60] and 1D Ta_4SiTe_4 samples.[61]

At 200K, the power factors of $Ni_{60}Fe_{40}$ and $Ni_{55}Fe_{45}$ are larger than those of the hot-pressed $YbAl_3$ sample[58] and are much larger than that of the Cu-Ni alloy[17]. However, the power factor values decrease rapidly with increasing Fe due to the increase in resistivity and with Ni concentration due to the reduction in the Seebeck coefficient. Since the TE power factor is our primary focus, the Ni-Fe composition range is restricted to 45% to 70% atomic Ni. In this range, the thermal conductivity generally increases with Ni content.

The effective thermal conductivity (κ_{eff}) of the $Ni_{60}Fe_{40}$ sample exhibits the best balance between power factor and thermal conductivity. As shown in Figure 7e κ_{eff} of the Ni-Fe samples under a 1K temperature gradient is 2 to 3 times higher than the κ_{eff} of pure Fe or Ni[11,62] in the above 200K range. Above room temperatures, κ_{eff} of the $Ni_{60}Fe_{40}$ alloy is still higher than that of pure copper, and previous studies of high power factor Cu-Ni alloys[63,64], reaching 600 $W/m.K$ for both $Ni_{60}Fe_{40}$ and $Ni_{70}Fe_{30}$ alloys.

As indicated by SEM, in the arc-melted samples, grains are significantly larger than the typical electron and phonon mean free paths in metals[65–67], eliminating the possibility of grain size influencing TE properties, especially the Seebeck coefficient. However, given the limited studies on Ni-Fe alloys as TE materials, further investigations with improved parameter control are essential to elucidate the role of microstructure in the thermoelectric performance of these alloys.

3. Conclusion

In summary, we built a database of binary metallic alloys and identified Ni-Fe as a potential candidate for active cooling applications. We used ML algorithms to identify the best molar

fraction corresponding to the largest Seebeck values in the 45%-55% Ni range. We then proceeded with experimental validation of this prediction. The highest Seebeck values were observed in $Ni_{55}Fe_{45}$ and $Ni_{45}Fe_{55}$ samples, consistent with ML prediction. The power factor and effective thermal conductivity of arc-melted Ni-Fe alloys with 45 to 70 atomic percent nickel were investigated over the 50K to 400K temperature range. Notably, the $Ni_{55}Fe_{45}$ and $Ni_{60}Fe_{40}$ alloys demonstrated a large peak power factor of 120 μW/cm·K² at 200 K. This metallic binary alloy is stable and is composed of cost-effective and abundant elements. The power factor value reported is one of the largest values reported in this temperature range. The effective thermal conductivity, κ_{eff}, at a 1K temperature difference was also calculated using the measured values of passive thermal conductivity and TE power factor. The largest κ_{eff} values exceeding 600 W/K·m at 400K were observed for $Ni_{60}Fe_{40}$ and $Ni_{70}Fe_{30}$ alloys, outperforming pure copper, Ni, Fe, and state-of-the-art Cu-Ni alloys under the same conditions. The microstructure of the arc-melted Ni-Fe ingots was characterized using SEM and EDS, providing insights into grain size and elemental distribution. The abnormal composition dependence of the absolute Seebeck coefficient at intermediate temperatures (200 K- 400 K) was also noted. A hypothesis suggesting that local concentration fluctuations account for this anomaly was tested using EDS analysis, which invalidated this explanation. Further research is needed to assess the effects of grain size, magnetic domains, and defects on the thermoelectric performance of Ni-Fe alloys. This study reveals the overlooked potential of Ni-Fe alloys for high-power factor applications, highlights the promise of magnetic transition metal alloys in the search for high-power factor metallic materials, and encourages further research into metallic thermoelectric materials for active cooling.

4. Methods

4.1. Selection and Optimization Details of Machine Learning:

Dataset:

Our initial dataset was obtained from the Landolt-Börnstein database[23] and a recent publication on Au-Ni alloys.[68] It comprises experimental Seebeck coefficient values for various binary metallic systems, recorded in wide temperature and atomic concentration ranges. The data was extracted and digitized manually using the GRABIT MATLAB tool, resulting in a total of 12,332 data points with 3,103 unique solid solutions. The Seebeck coefficient values span from 40 to -85 μV/K, covering a temperature range of 0 to 1500 K as depicted in Figure 2 and Figure S1(a), respectively. [9]

Figure 2(a) demonstrates the distribution of the Seebeck coefficient in this dataset, revealing that most metals have small Seebeck values (a few μV/K), which is one of the main reasons for the TE community to focus on semiconductors instead. Figure S1(b) provides an overview of the metals included in the dataset, showing that Ni is the most frequently used material. It is followed by Pd, Cr, Pt, and Cu. This information is presented in an alternative format in Figure 2b.

4.2. Feature Selection

To build generalizable data-driven models, it is important to include features that not only capture the trend of Seebeck coefficients across different metallic alloys but also uniquely represent them. In this regard, a Composition-Based Feature Vector (CBFV)[69] technique was used to derive features from the chemical formula, utilizing the Materials Agnostic Platform for Informatics and Exploration (Magpie). [70] Furthermore, temperature and crystallinity information (Single-crystal /Polycrystalline)by level encoding (1/0) are also added to the feature list, which results in a total of 156 input features, and the Seebeck coefficient as the target value. A detailed table summarizing the input features is provided in Table S1. These features are commonly used in the field of materials informatics of TE.[71–75] The correlation analysis, as shown in Figure S3, indicates that many features exhibit strong statistical correlation. In most cases, it is advisable to remove one of two highly correlated features since they convey redundant information. Such correlations can hinder model convergence, degrade predictive performance, and affect interpretability.

The dimensionality of the input features was reduced by applying a correlation coefficient threshold of 0.5.[76,77] This means that only those features with an absolute correlation coefficient less than 0.5 with other features were kept for ML model building.

This process reduced the number of input features to 19, which were used to predict the Seebeck values.

4.3. Training and Testing of the Models:

To avoid any perceived bias during training, we employed a data-driven approach for splitting the dataset. We performed K-means[78] clustering analysis (using Euclidean distance as the similarity metric) on the dataset, and the Silhouette score, as shown in Figure S4, suggests that the dataset contains two distinct clusters. Each cluster represents a different group of data points that share common characteristics. Cluster 1 consists of 8,743 data points, while Cluster 2 contains 3,589 data points. To ensure that the models learn from both types of data distributions, we randomly selected 70% of the data from each cluster to form the training set (8,632 data points), with the remaining 30% used as the testing set (3,700 data points). This approach ensures a more balanced representation of samples from both clusters in the training and testing sets.

4.4. Hyperparameter Tuning:

Hyperparameter tuning helps improve model performance and prevents overfitting. To optimize the hyperparameters, we used BayesSearchCV from the scikit-optimize library[79] in Python. BayesSearchCV employs a Gaussian Process Regression as a surrogate model for hyperparameter optimization. An acquisition function is used to determine which hyperparameter combinations to evaluate next, with Expected Improvement (EI) as the default acquisition function. The EI function estimates the expected improvement over the current best result.[80] During optimization, 10-fold cross-validation from

the scikit-learn library [81] was applied, which partitioned the training set into ten subsets. Each model is trained in nine subsets and validated on the remaining subset. The testing set remained unseen by the models during cross-validation. The lower and upper boundaries for the hyperparameters are provided in Table S2. The number of iterations for the optimization process was set to 50. The optimized hyperparameters for each model in each case of splitting are shown in Table S2. The performance of these models was evaluated by comparing two key metrics, namely the coefficient of determination (R^2) and mean squared error (MSE).

4.5. Experimental Methods

Iron powder with 99.5% purity and nickel powder with 99.996% purity were weighed to 5 grams per sample and mixed in an argon-filled glovebox. The powder mixtures were then hot-pressed into solid bulk samples at 800^{o}C under 56 MPa pressure for 300 seconds using an OTF-1700X-RHP4 hot-press setup from MTI Corporation. The solid bulk samples were later arc-melted in Ar protected chamber to form a Ni-Fe solid solution. Each sample was melted and flipped twice for homogeneity. Then, it was melted without flipping, allowing bubbles and voids to diffuse to the top of the sample, which was then cut out. The central-bottom part of the arc-melted samples was then sectioned into approximately 2mm×2mm×10mm bar shape. Transport properties were measured using the Thermal Transport Option of Quantum Design PPMS Versalab. A heater was attached to one side of the sample to create a 3% rise in temperature. The other side was connected to a heat sink. The resulting voltage difference and temperature difference under steady state were measured along the length of the sample to extract the Seebeck coefficient and the thermal conductivity. The heater and the heat sink contact (copper coated with gold) were then used to send current along the sample. The voltage was measured using side probes, enabling 4-probe electrical conductivity measurements. The XRD characterization is performed using an Empyrean X-ray diffractometer from Malvern-Panalytical on the sectioned as-arc-melted ingots. SEM/EDS is performed on an FEI Quanta 650 Scanning Electron Microscope (SEM).

Author contributions

Shuai Li: Methodology and formal analysis (sample preparation, XRD, transport measurements), writing-original draft; Sree Sourav Das: Methodology, data curation, and formal analysis (database, ML), writing-original draft; Haobo Wang: Methodology (sample preparation, SEM), Sujit Bati: Methodology (DFT calculation), Prasanna Balachandran: Formal analysis (ML algorithms), writing-review & editing; Junichiro Shiomi Formal analysis (ML algorithms), writing-review & editing; Jerry Floro: Formal analysis, supervision, writing-review & editing, Mona Zebarjadi, conceptualization, supervision, writing-original draft, Funding acquisition, Kacper Pryga: Methodology (KKR-CPA and transport calculation), Bartlomiej Wiendlocha: Formal analysis, supervision (KKR-CPA and transport calculation)

Conflicts of interest

There are no conflicts to declare.

Data availability

Data for this article, raw SEM and EDS scans are available at https://doi.org/10.18130/V3/AOSTG4. The dataset containing Seebeck coefficients for binary metallic systems will be available soon **in** figshare data repository.

Acknowledgments

This work is supported by NSF grant number 2421213. MZ and SSD acknowledge discussions with Masato Onishi and Ryo Yoshida on the details of the ML algorithm. SSD and SB acknowledge the Rivanna cluster of UVA used for the computational part. KP and BW were supported by the 'Excellence Initiative – Research University' program at AGH University of Krakow and acknowledge Polish high-performance computing infrastructure PLGrid (HPC Center: ACK Cyfronet AGH) for providing computer facilities and support within computational grant no. PLG/2024/017306.

Notes and references

1 H. J. Goldsmid, *Springer Series in Materials Science*, 2016, **121**, 153–195.

2 N. Jaziri, A. Boughamoura, J. Müller, B. Mezghani, F. Tounsi and M. Ismail, *Energy Reports*, 2020, **6**, 264–287.

3 Z. Soleimani, S. Zoras, B. Ceranic, S. Shahzad and Y. Cui, *Sustainable Energy Technologies and Assessments*, 2020, **37**, 100604.

4 D. L. Medlin and G. J. Snyder, *Curr Opin Colloid Interface Sci*, 2009, **14**, 226–235.

5 G. Chen, M. S. Dresselhaus, G. Dresselhaus, J. P. Fleurial and T. Caillat, *International Materials Reviews*, 2003, **48**, 45–66.

6 M. J. Adams, M. Verosky, M. Zebarjadi and J. P. Heremans, *Phys Rev Appl*, 2019, **11**, 054008.

7 M. J. Adams, M. Verosky, M. Zebarjadi and J. P. Heremans, *Int J Heat Mass Transf*, DOI:10.1016/j.ijheatmasstransfer.2018.12.154.

8 S. Li, K. Snyder, M. S. Akhanda, R. Martukanitz, M. Mitra, J. Poon and M. Zebarjadi, *Int J Heat Mass Transf*, 2022, **195**, 123181.

9 M. Zebarjadi, *Appl Phys Lett*, 2015, **106**, 203506.

10 H. Zhu, J. Mao, Y. Li, J. Sun, Y. Wang, Q. Zhu, G. Li, Q. Song, J. Zhou, Y. Fu, R. He, T. Tong, Z. Liu, W. Ren, L. You, Z. Wang, J. Luo, A. Sotnikov, J. Bao, K. Nielsch, G. Chen, D. J. Singh and Z. Ren, *Nat Commun*, 2019, **10**, 270.

11 C. L. Foiles, *Electrical Resistivity, Thermoelectrical Power and Optical Properties*, Springer-Verlag, 1985, vol. 15b.

12 K. Vandaele, S. J. Watzman, B. Flebus, A. Prakash, Y. Zheng, S. R. Boona and J. P. Heremans, *Materials Today Physics*, 2017, **1**, 39–49.

13 W. Fulkerson, J. P. Moore and D. L. McElroy, *J Appl Phys*, 1966, **37**, 2639–2653.

14 D. M. Rowe, Ed., *CRC Handbook of Thermoelectrics*, CRC Press, Boca Raton, 2018.

15 F. Garmroudi, M. Parzer, A. Riss, C. Bourgès, S. Khmelevskyi, T. Mori, E. Bauer and A. Pustogow, *Sci Adv*, DOI:10.1126/SCIADV.ADJ1611/SUPPL_FILE/SCIADV.ADJ1611_SM.PDF.

16 S. J. Watzman, R. A. Duine, Y. Tserkovnyak, S. R. Boona, H. Jin, A. Prakash, Y. Zheng and J. P. Heremans, *Phys Rev B*, 2016, **94**, 144407.

17 J. Mao, Y. Wang, H. S. Kim, Z. Liu, U. Saparamadu, F. Tian, K. Dahal, J. Sun, S. Chen, W. Liu and Z. Ren, *Nano Energy*, 2015, **17**, 279–289.

18 T. Maeda and T. Somura, *J. Phys. Soc. Jpn.*, 1978, **44**, 148–153.

19 Y. Tanji, H. Moriya and Y. Nakagawa, *https://doi.org/10.1143/JPSJ.45.1244*, 1978, **45**, 1244–1248.

20 A. Giannuzzi, H. Tomaschke and K. Schröder, *Philosophical Magazine*, 1970, **21**, 479–493.

21 C. Y. Ho, R. H. Bogaard, T. C. Chi, T. N. Havill and H. M. James, *Thermochim Acta*, 1993, **218**, 29–56.

22 K. Schröder, M. J. Yessik and N. P. Baum, *J Appl Phys*, 1966, **37**, 1019–1021.

23 Landolt Bornstein, http://www.springermaterials.com/docs/index.html.

24 K. V. Selvan and M. S. Mohamed Ali, *IEEE Trans Electron Devices*, 2018, **65**, 3394–3400.

25 C. W. Yang, D. B. Williams and J. I. Goldstein, *Journal of Phase Equilibria*, 1996, **17**, 522–531.

26 W. Xiong, H. Zhang, L. Vitos and M. Selleby, *Acta Mater*, 2011, **59**, 521–530.

27 R. A. Howald, *Metall Mater Trans A Phys Metall Mater Sci*, 2003, **34 A**, 1759–1769.

28 T. Komabayashi, K. Hirose and Y. Ohishi, DOI:10.1007/s00269-012-0490-3.

29 C. H. Xia, Y. Wang, J. J. Wang, X. G. Lu and L. Zhang, *J Alloys Compd*, 2021, **853**, 157165.

30 K. Li, C.-C. Fu, M. Nastar and F. Soisson, *Phys Rev B*, 2023, **107**, 94103.

31 G. Cacciamani, J. De Keyzer, R. Ferro, U. E. Klotz, J. Lacaze and P. Wollants, *Intermetallics (Barking)*, 2006, **14**, 1312–1325.

32 D. Ma, Z. Wang, H. Zhao, J. Chen and F. Ke, *Ferroelectrics*, 2021, **571**, 175–182.

33 P. Bag, Y. C. Su, Y. K. Kuo, Y. C. Lai and S. K. Wu, *Phys Rev Mater*, 2021, **5**, 085003.

34 Y. Mokrousov, H. Zhang, F. Freimuth, T. Farrellt and D. Greig, *Journal of Physics C: Solid State Physics*, 1970, **3**, 138.

35 R. Tibshirani, *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996, **58**, 267–288.

36 T. Chen and C. Guestrin, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, New York, NY, USA, 2016, pp. 785–794.

37 L. Breiman, *Mach Learn*, 2001, **45**, 5–32.

38 J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua and A. Lopez, *Neurocomputing*, 2020, **408**, 189–215.

39 A. A. Freitas, in *Machine Learning and Knowledge Extraction: Third IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9 International Cross-Domain Conference, CD-MAKE 2019, Canterbury, UK, August 26–29, 2019, Proceedings*, Springer-Verlag, Berlin, Heidelberg, 2019, pp. 48–66.

40 C. J. Hickson and S. J. Juras, *Can Mineral*, 1986, **24**, 585–589.

41 A. Seijas Da Silva, A. Hartert, V. Oestreicher, J. Romero, E. Coronado, V. Lloret Segura and G. Abellán, *ECS Meeting Abstracts*, 2023, **MA2023-01**, 2266.

42 C. Liang, W. Pan, P. Zou, P. Liu, K. Liu, G. Zhao, H. J. Fan and C. Yang, *Small*, 2022, **18**, 2203663.

43 H. Wijn, *3d, 4d and 5d Elements, Alloys and Compounds*, Springer-Verlag, Berlin/Heidelberg, 1997, vol. 32A.

44 M. Hayase, M. Shiga and Y. Nakamura, *J Physical Soc Japan*, 1973, **34**, 925–931.

45 H. Ebert, D. Ködderitzsch and J. Minár, *Reports on Progress in Physics*, 2011, **74**, 096501.

46 H. Ebert et al, The Munich SPR-KKR package, version 8.6, https://www.ebert.cup.uni-muenchen.de/index.php/de/software/13-sprkkr, (accessed 25 May 2025).

47 R. Kubo, *https://doi.org/10.1143/JPSJ.12.570*, 2013, **12**, 570–586.

48 D. A. Greenwood, *Proceedings of the Physical Society*, 1958, **71**, 585.

49 W. H. Butler, *Phys Rev B*, 1985, **31**, 3260.

50 J. C. Swihart, W. H. Butler, G. M. Stocks, D. M. Nicholson and R. C. Ward, *Phys Rev Lett*, 1986, **57**, 1181.

51 Q. Wei, S. A. Gilder and B. Maier, *Phys Rev B Condens Matter Mater Phys*, 2014, **90**, 144425.

52 K. Pryga and B. Wiendlocha, *To be Published*.

53 B. Wiendlocha, *Phys Rev B*, 2018, **97**, 205203.

54 A. Vernes, H. Ebert and J. Banhart, *Phys Rev B*, 2003, **68**, 134404.

55 J. Banhart and H. Ebert, *Solid State Commun*, 1995, **94**, 445–449.

56 C. Y. Ho, M. W. Ackerman, K. Y. Wu, T. N. Havill, R. H. Bogaard, R. A. Matula, S. G. Oh and H. M. James, *J Phys Chem Ref Data*, 1983, **12**, 183–322.

57 C. L. Chen, T. H. Wang, Z. G. Yu, Y. Hutabalian, R. K. Vankayala, C. C. Chen, W. P. Hsieh, H. T. Jeng, D. H. Wei and Y. Y. Chen, *Advanced Science*, 2022, **9**, 2201353.

58 X. Zhang, H. Wang, W. Cui, X. Xie, P. Zhai, D. He and W. Zhao, *Materials Today Physics*, 2023, **32**, 101008.

59 D. M. Rowe, G. Min and V. L. Kuznestsov, *Philos Mag Lett*, 1998, **77**, 105–108.

60 S. Shimizu, J. Shiogai, N. Takemori, S. Sakai, H. Ikeda, R. Arita, T. Nojima, A. Tsukazaki and Y. Iwasa, *Nat Commun*, 2019, **10**, 1–7.

61 T. Inohara, Y. Okamoto, Y. Yamakawa, A. Yamakage and K. Takenaka, *Appl Phys Lett*, 2017, **110**, 183901.

62 G. K. White, *Thermal Conductivity of Pure Metals and Alloys*, Springer-Verlag, 1991, vol. 15c.

63 J. Mao, Y. Wang, H. S. Kim, Z. Liu, U. Saparamadu, F. Tian, K. Dahal, J. Sun, S. Chen, W. Liu and Z. Ren, *Nano Energy*, 2015, **17**, 279–289.

64 S. Li, K. Snyder, M. S. Akhanda, R. Martukanitz, M. Mitra, J. Poon and M. Zebarjadi, *Int J Heat Mass Transf*, 2022, **195**, 123181.

65 S. Sæther, M. F. Erichsen, S. Xiao, Z. Zhang, A. Lervik and J. He, *AIP Adv*, 2022, **12**, 65301.

66 D. Gall, *J Appl Phys*, 2016, **119**, 85101.

67 A. Jain and A. J. H. McGaughey, *Phys Rev B*, 2016, **93**, 081206.

68 F. Garmroudi, M. Parzer, A. Riss, C. Bourgès, S. Khmelevskyi, T. Mori, E. Bauer and A. Pustogow, *Sci Adv*, DOI:10.1126/SCIADV.ADJ1611.

69 R. J. Murdock, S. K. Kauwe, · Anthony, Y.-T. Wang, · Taylor and D. Sparks, DOI:10.26434/CHEMRXIV.11879193.V1.

70 L. Ward, A. Agrawal, A. Choudhary and C. Wolverton, *npj Computational Materials 2016 2:1*, 2016, **2**, 1–7.

71 H. M. Yuan, S. H. Han, R. Hu, W. Y. Jiao, M. K. Li, H. J. Liu and Y. Fang, *Materials Today Physics*, 2022, **25**, 100706.

72 X. Jia, Y. Deng, X. Bao, H. Yao, S. Li, Z. Li, C. Chen, X. Wang, J. Mao, F. Cao, J. Sui, J. Wu, C. Wang, Q. Zhang and X. Liu, *NPJ Comput Mater*, 2022, **8**, 34.

73 Y. Liu, T. Zhao, W. Ju and S. Shi, *Journal of Materiomics*, 2017, **3**, 159–177.

74 N. Parse, C. Pongkitivanichkul and S. Pinitsoontorn, *Energies (Basel)*, DOI:10.3390/en15030779.

75 L. M. Antunes, Vikram, J. J. Plata, A. V Powell, K. T. Butler and R. Grau-Crespo, in *Machine Learning in Materials Informatics: Methods and Applications*, American Chemical Society, 2022, vol. 1416, p. 1.

76 N. K. Barua, E. Hall, Y. Cheng, A. O. Oliynyk and H. Kleinke, *Chemistry of Materials*, 2024, **36**, 7089–7100.

77 Y. Liu, Z. Mu, P. Hong, Y. Yang and C. Lin, *Nanoscale*, 2025, **17**, 2200–2214.

78 S. Lloyd, *IEEE Trans Inf Theory*, 1982, **28**, 129–137.

79 T. Head, M. Kumar, H. Nahrstaedt, G. Louppe and I. Shcherbatyi, *Zenodo*, 2020, preprint, DOI: 10.5281/zenodo.4014775.

80 V. Nguyen, S. Gupta, S. Rana, C. Li and S. Venkatesh, in *Asian Conference on Machine Learning*, 2017.

81 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and É. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.

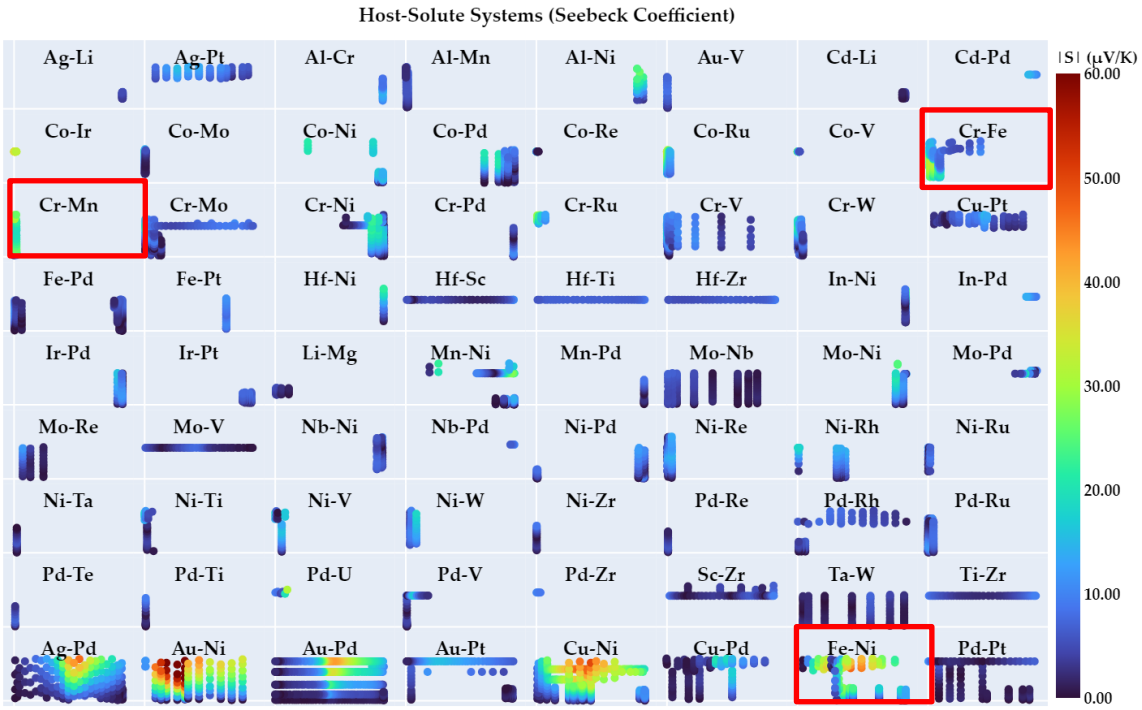


Figure 1. Visual representation of part of our database. Seebeck values of less than a few microvolts per Kelvin are eliminated and the rest are represented here. Each subplot represents the host-solvent binary alloy labelled in the plot. For example, Cu-Ni means Cu is the host, Ni is the solvent. Each subplot has an x-axis scale of 0 to 1 representing the molar fraction of the solvent. The y-axis is the measurement temperature from 0-400 K for all subplots. The colour represents the absolute value of the Seebeck coefficient. Promising yet less-explored candidates are highlighted with red rectangular boxes.

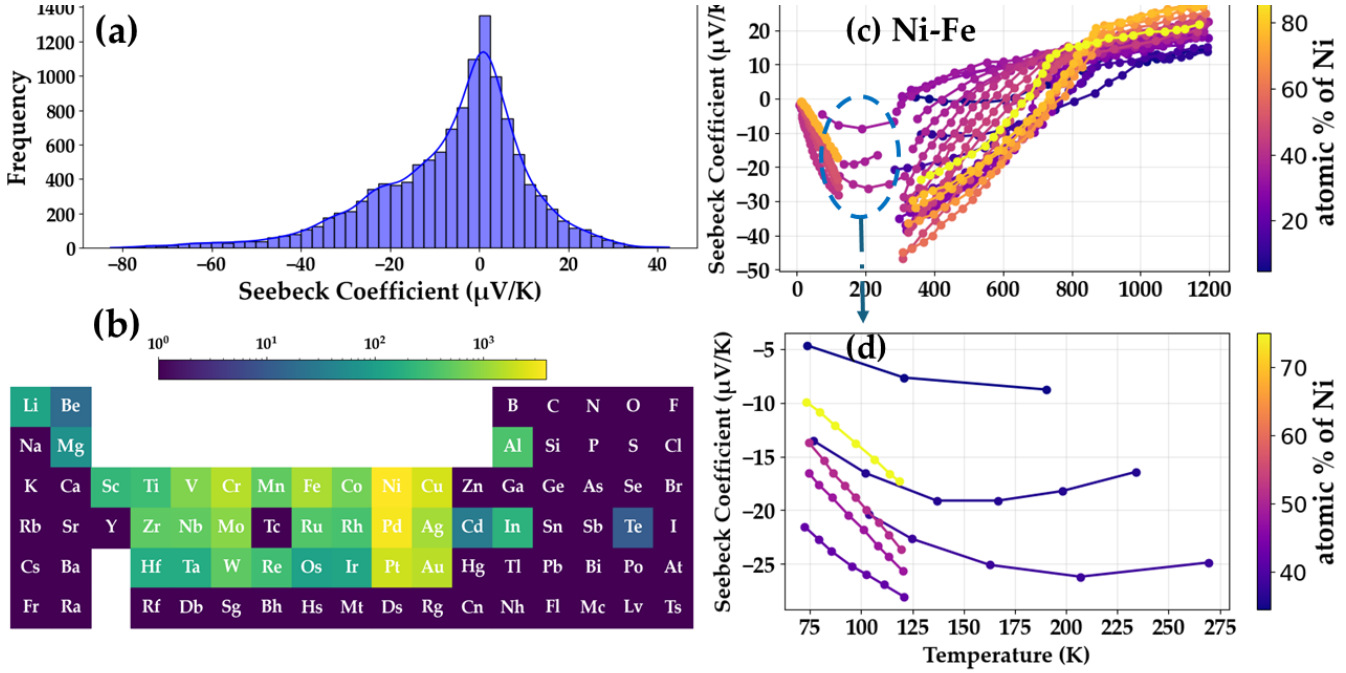


Figure 2. (a) Distribution of Seebeck coefficient in the database, (b) Frequency of the elements in the dataset, for example, Ni and Pd are the most frquently used metals, and (c) Temperature dependent Seebeck coefficients of Fe-Ni system upto 1200K with respect to different atomic % of Ni, (d) highlighting low temperature region from 50 K to 300 K.

ARTICLE

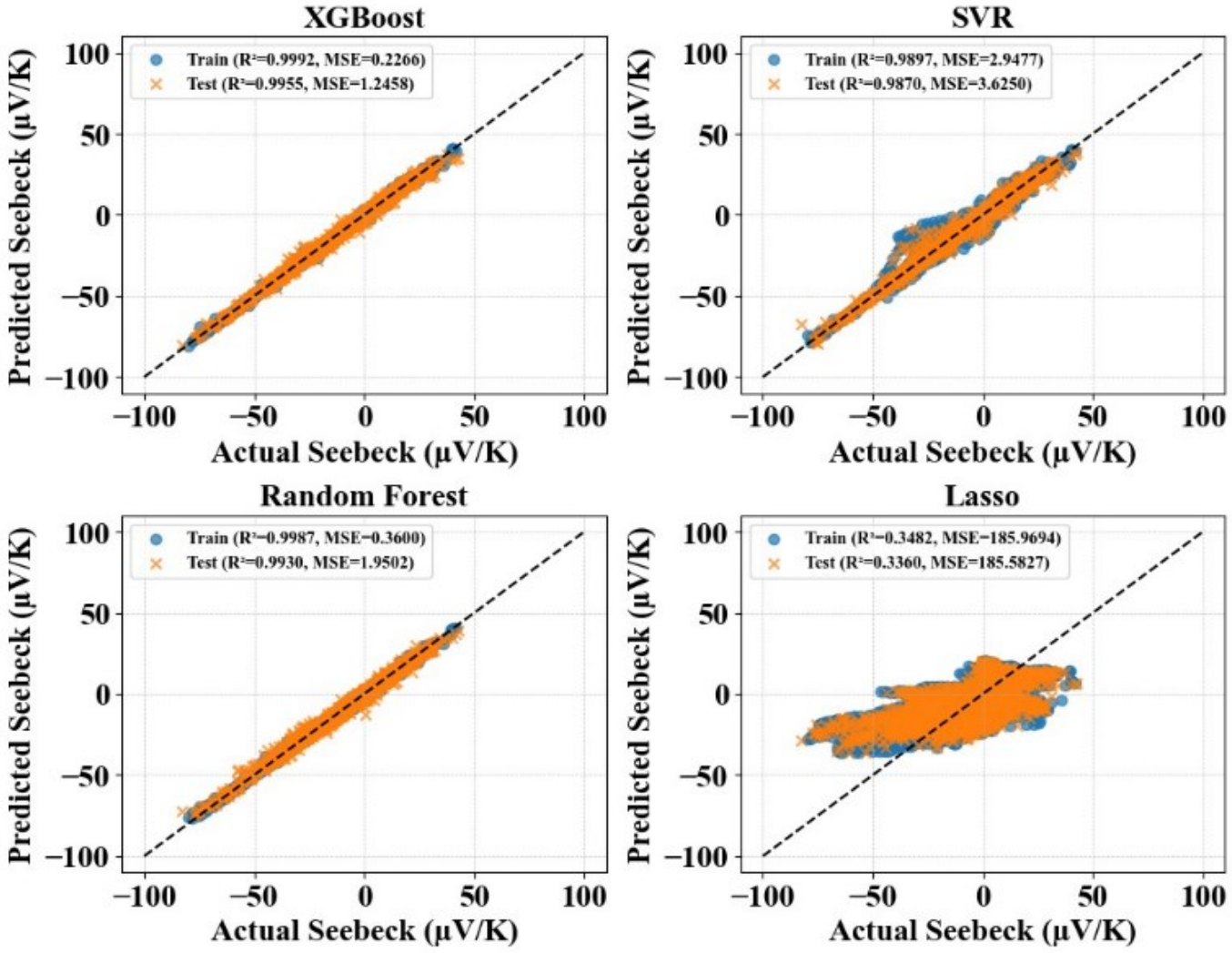


Figure 3. Predicted vs Actual Seebeck coefficients from XGB, SVR, RF and Lasso. Tree and Kernel-based models outperformed the linear model

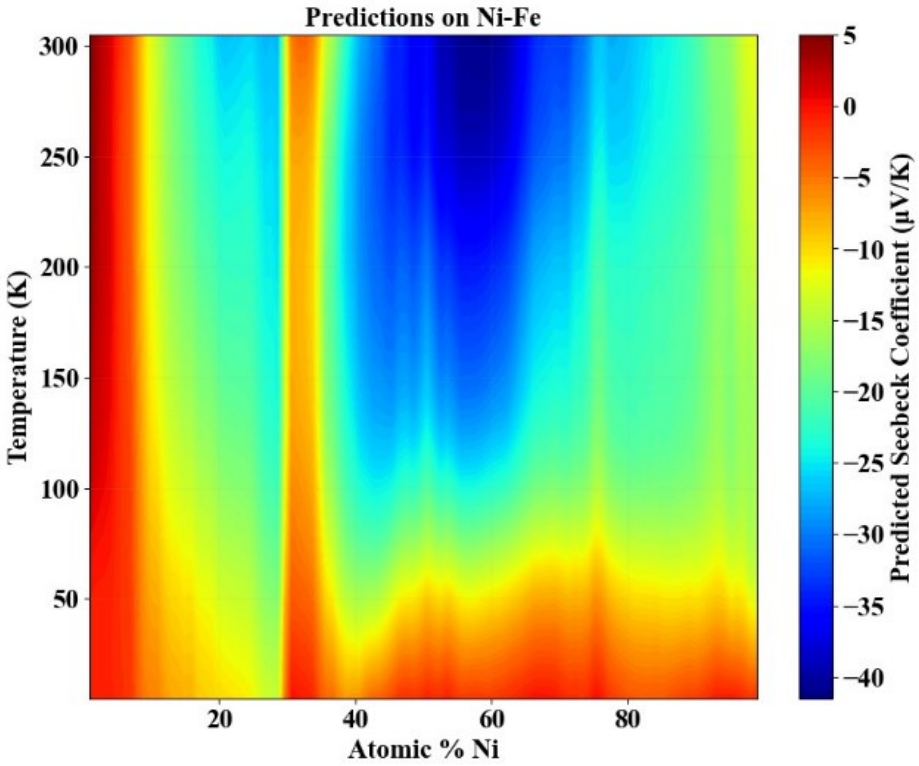


Figure 4 Heat map of Ni-Fe alloy averaging over all predictions from all optimized models.

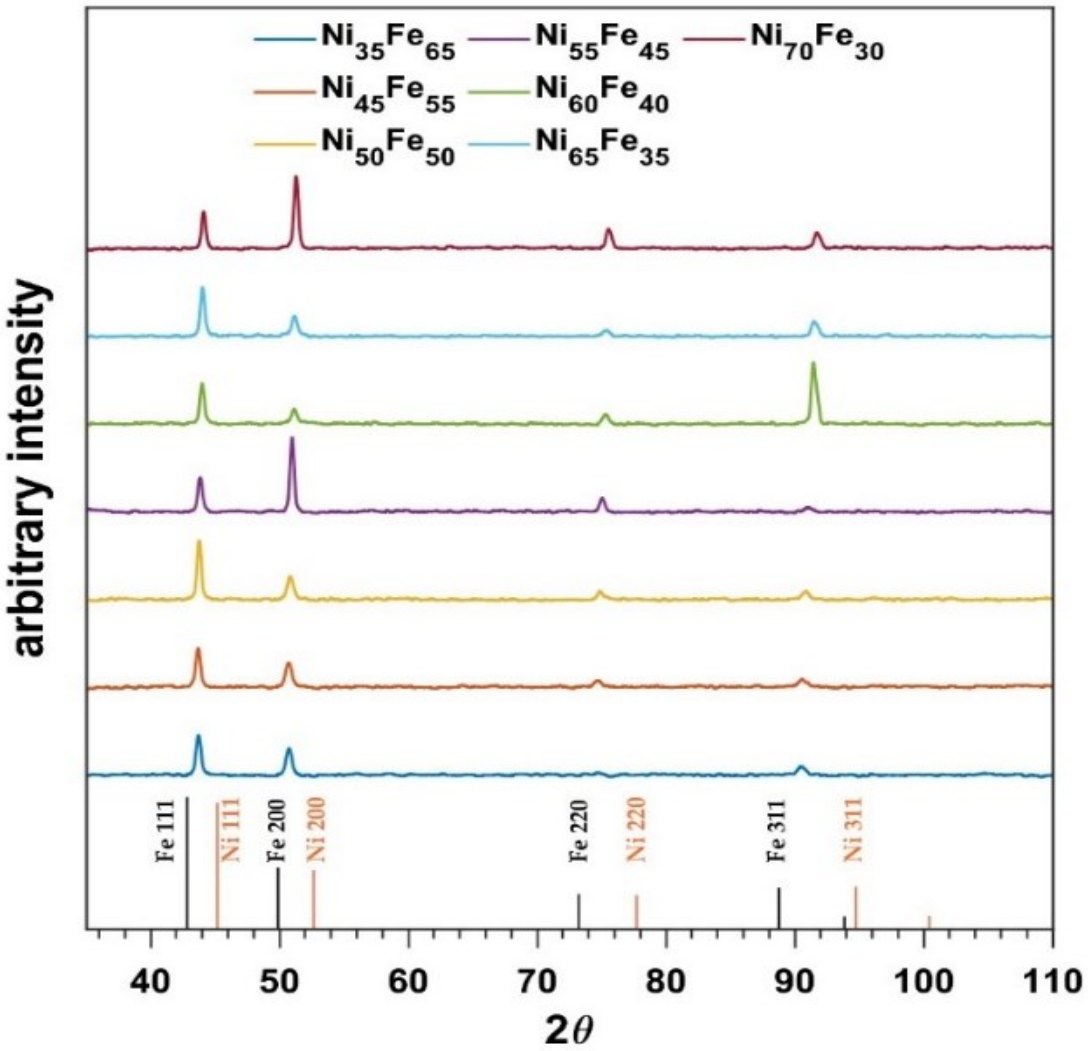
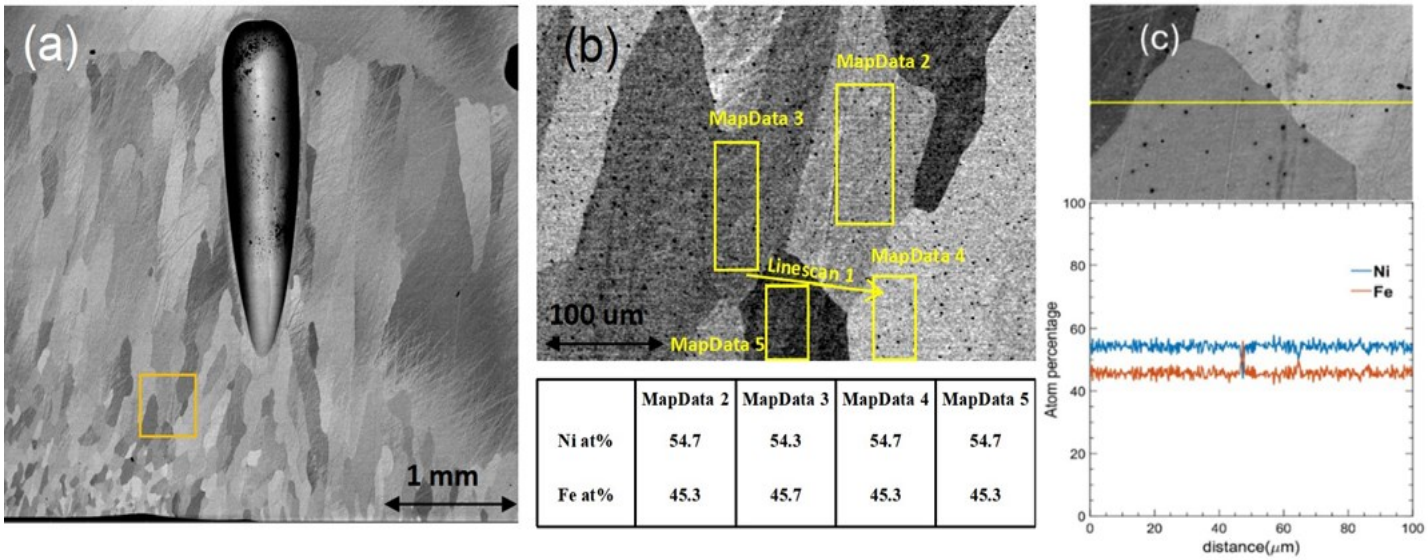


Figure 5. XRD results of arc-melted Ni-Fe samples. The black and orange straight lines on the bottom are FCC iron and nickel reference peaks.



	MapData 2	MapData 3	MapData 4	MapData 5
Ni at%	54.7	54.3	54.7	54.7
Fe at%	45.3	45.7	45.3	45.3

Figure 6. SEM backscattered electron image on the cross-section of the arc-melted $Ni_{55}Fe_{45}$ sample(a), the yellow square marked area is the chosen EDS mapping area. This area is enlarged in (b) where the EDS mapping locations on different grains are reported and summarized in the table. (c) EDS line scan across the grains over the line scan 1 shown in (b)

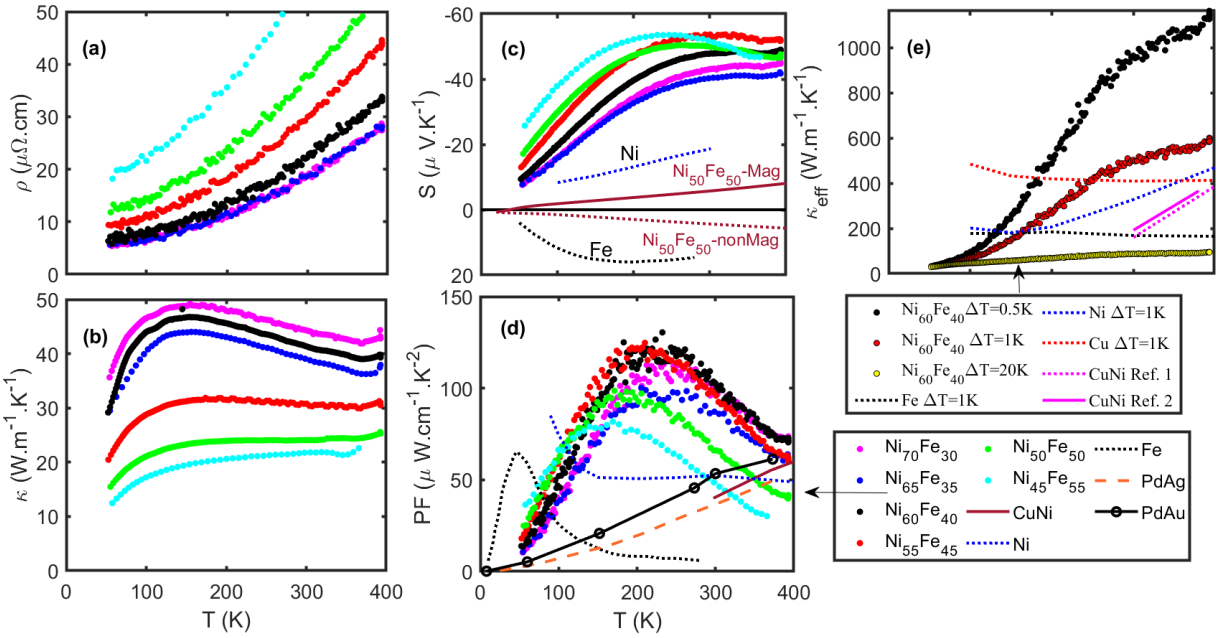


Figure 7. (a) electrical resistivity, (b) thermal conductivity, (c) Seebeck coefficient, (d) power factor, (e) effective thermal conductivity of Ni-Fe samples. The lower Legend on the right bottom corner is shared for a-d. Ni and Fe reference data in (c) are taken from Ref[11] . The calculated Seebeck coefficient for 50% Ni-Fe (KKR-CPA) is also shown, the non-magnetic calculations fail to predict the correct sign of the Seebeck coefficient. The magnetic calculations predicts the sign correctly, however, the absolute values are much smaller. The difference is contributed by additional scattering processes, most likely associated with the magnon-drag. The power factor of Ni, Fe[12], CuNi[17], PdAg and PdAu[21,56] are compared with that of NiFe alloy in d. Effective thermal conductivity of Ni-Fe alloys compared with pure Cu, pure Fe, pure Ni[62], and Cu-Ni alloys (labeled as Ref 1 and 2)[42,43].