Research Article

Domain knowledge-guided interpretive machine learning: formula discovery for the oxidation behavior of ferritic-martensitic steels in supercritical water

Bin Cao[1] , **Shuang Yang[1]**, **Ankang Sun[1]**, **Ziqiang Dong[1,*]**, **Tong-Yi Zhang[1,2,*]**

[1]Materials Genome Institute, Shanghai University, Shanghai 200444, China.
[2]Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511400, Guangdong, China.

*Correspondence to: Prof. Tong-Yi Zhang, Materials Genome Institute, Shanghai University, 333 Nanchen Road, Shanghai 200444, China. E-mail: zhangty@shu.edu.cn; Prof. Ziqiang Dong, Materials Genome Institute, Shanghai University, 333 Nanchen Road, Shanghai 200444, China. E-mail: zqdong@shu.edu.cn

Abstract

A general formula with high generalization and accurate prediction power is highly desirable for science, technology and engineering. In addition to human beings, artificial intelligence algorithms show great promise for the discovery of formulas. In this study, we propose a domain knowledge-guided interpretive machine learning strategy and demonstrate it by studying the oxidation behavior of ferritic-martensitic steels in supercritical water. The oxidation Cr equivalent is, for the first time, proposed in the present work to represent all contributions of alloying elements to oxidation, derived by our domain knowledge and interpretive machine learning algorithms. An open-source tree classifier for linear regression algorithm is also, for the first time, developed to materialize the formula with collected data. This algorithm effectively captures the linear correlation between compositions, testing environments and oxidation behaviors from the data. The sure independence screening and sparsifying operator algorithm finally assembles the information derived from the tree classifier for linear regression algorithm, resulting in a general formula. The general formula with the determined parameters has the power to predict, quantitatively and accurately, the oxidation behavior of ferritic-martensitic steels with multiple alloying elements exposed to various supercritical water environments, thereby providing guidance for the design of anti-oxidation steels and hence promoting the development of power plants with improved safety. The present work demonstrates the power of domain knowledge-guided interpretive machine learning with respect to the data-driven discovery of physics-informed formulas and the acceleration of materials informatics development.

INTRODUCTION

The rapid development of materials informatics[1–5], artificial intelligence (AI) and machine learning (ML) techniques has led to a new paradigm of data-driven discovery of novel materials, state-of-the-art experimental and computational methods and scientific laws and formulas. The number of publications on materials informatics has increased exponentially in the past decade and materials informatics has achieved great success in many areas[6–10]. For example, Xue *et al.*[11] proposed an adaptive design iteration strategy by tightly coupling ML with experiments, which sequentially identifies the next experiments by using efficient global optimization to balance the trade-off between exploitation and exploration. This adaptive design strategy, also known as active learning, starts from an initial dataset of 22 alloys, runs nine feedback loops in the search space of ~800,000 alloys and successfully finds very low thermal hysteresis NiTi-based shape memory alloys. With a dataset of electronic density of states (DOS) generated by high-throughput density functional theory-based computations, Fung *et al.*[12] used a convolutional neural network-based ML model to automatically obtain the key features for the accurate prediction of catalytic properties, such as adsorption energies. The ML model yields a DOSnet, which has the capacity to provide physically meaningful predictions and insights by predicting responses to external perturbations to the electronic structure without additional calculations.

Attia *et al.*[13] developed an ML methodology to efficiently optimize a parameter space specifying the current and voltage profiles of six-step, 10-min fast-charging protocols for maximizing battery cycle life. They trained an elastic net ML model to predict battery charging/discharging life using data only from the first few cycles and employed a Bayesian optimization algorithm to reduce the number of experiments by balancing exploration and exploitation to efficiently probe the parameter space of charging protocols. With such an approach, they identified and validated high-cycle-life charging protocols among 224 candidates in 16 days. Saito *et al.*[14] conducted an image process by using U-Net based on a convolutional encoder-decoder network to segment and identify the thickness of atomic layer flakes from optical microscopy images, achieving a success rate of 70–80% in distinguishing monolayer and bilayer MoS_2 and graphene.

ML is achieving remarkable success in materials science and engineering[15,16] and will achieve even greater success if it can become more transparent and interpretive. Theoretically, AI and ML are based on statistics without utilizing any other scientific laws, principles and (physical) equations and most AI and ML algorithms perform as "black-box" systems[17–23]. Considerable efforts, such as physics-informed neural networks[24], symbolic regression and Shapley additive explanations (SHAP)[25], are being carried out to enhance the interpretability of ML models. Obviously, significant further endeavors are required to make ML models interpretive. The strategy proposed in the present work, i.e., domain knowledge-guided interpretive ML, might pave the way for the discovery of mathematical formulas.

In the present work, we propose a domain knowledge-guided interpretive ML strategy to make ML models interpretable and have more physical sense and apply this strategy to the data-driven discovery of formulas regarding the oxidation of ferritic-martensitic (FM) steels in supercritical water (SCW). Although the use of SCW in power plants can achieve enhanced thermal efficiency with simplified plant design and improved safety, it requires high anti-oxidation materials because SCW is a strong oxidant[26] beyond the supercritical point (at 374 °C and 22.1 MPa). FM steels are some of the most promising structural materials for use in SCW-cooled power plants, owing to their high elevated temperature strength, high creep resistance, high thermal conductivity, low swelling behavior under irradiation, low thermal expansion coefficient, and low susceptibility to stress oxidation cracking up to 600 °C[27,28].

The oxidation behavior of FM steels in SCW environments has been investigated extensively through experimental approaches[29–34]. The current understanding of the corrosion occurring in high-temperature water environments is associated with the chemistry and physics of the water (density and dielectric constant of the medium). In high-temperature water with a low density/dielectric constant (< 0.1 g/cm^3/ < 10), the direct oxidation reaction occurs between materials and oxidants (oxygen and/or water), while in high-temperature water with a high density/dielectric constant (> 0.1 g/cm^3/ > 10), electrochemical oxidation proceeds via the partial anodic and cathodic reactions[30]. The oxidation behavior of FM alloys in SCW depends on the alloy composition and oxidation environment. Ampornrat and Was[29] experimentally investigated the corrosion behavior of three FM alloys (T91, HCM12 A and HT-9) in SCW at temperatures ranging from 400 to 600 °C with dissolved oxygen concentrations of < 10 ppb (deaerated conditions), 100 ppb and 300 ppb. Cr is conducive to the formation of a continuous inner Cr-Fe oxide layer to block the transportation of Fe (outward) and oxygen (inward); therefore, a higher concentration of Cr generally corresponds to a lower oxidation rate of FM steels in SCW. Furthermore, Si facilitates the formability of Cr-rich areas in FM steels[35]. Dong *et al.*[36] reported that with increasing Mn content, the oxide scale becomes discontinuous and thicker, and thus Mn might be harmful to the oxidation of FM steels in SCW.

Significant progress has been achieved in the investigation and understanding of FM steel oxidation in various SCW environments[29,30,37–39], as evidenced by the Arrhenius equation of $\Delta W = k_{\text{eff}} \exp\left(-\frac{Q}{nRT}\right) t^{\frac{1}{n}}$ [29], where ΔW per unit area (mg/dm^2) denotes the oxidation weight gained, k_{eff} (mg/dm^2/h) is the effective oxidation rate constant, Q is the activation energy of oxidation (J/mol), R is the gas constant (8.314 J/mol·K), T is the absolute temperature (K), t is time (h) and $1/n$ is the exponent over time.

Oxidation is clearly a thermally activated process, and the associated thermodynamics and kinetics are greatly dependent on the material compositions and environmental variables. In experimental investigations, individual researchers adjust only one or a few experimental conditions and the obtained result and formula are valid only for the FM steels and SCW environments and periods tested. To the best of our knowledge, no generalized formula has been established for the description and/or prediction of the oxidation of FM steels with any given alloying elements exposed to various SCW environmental conditions. The present work adopts domain knowledge-guided interpretive ML to discover a generalized formula for the oxidation of FM steels in SCW, which will promote the development of green and safe power plants. In addition to exposure time, the investigated FM steels cover 11 alloying elements, and the studied SCW environments include temperature, dissolved oxygen concentration (DOC) and pressure.

Our domain knowledge of oxidation suggests a dimensionless Arrhenius equation of $\frac{\Delta w}{\Delta w_{00}} = \frac{\Delta w_0}{\Delta w_{00}} \exp\left(-\frac{Q}{RT}\right)$ $\left(\frac{t}{t_0}\right)^m g\left(\frac{(\text{DOC}+r)}{(\text{DOC})_0}\right)$, where $\Delta w_{00} = 1$ mg/dm^2 and is the reference weight gain, Δw_0 is a prefactor, $g\left(\frac{\text{DOC}+r}{(\text{DOC})_0}\right)$ is a function of the DOC, $(\text{DOC})_0 = 1$ ppb and is the reference DOC, r is a parameter introduced to account for the oxidation in deaerated SCW and m is the exponent of time. With this Arrhenius equation, the present work integrates interpretive ML algorithms, including SHAP[25], extreme gradient boosting (Xgboost)[40] and the sure independence screening and sparsifying operator (SISSO)[41], and more significantly, a newly developed classifier model, the tree classifier for linear regression (TCLR)[42], to discover a generalized formula from data for the oxidation of FM steels in SCW. Recently, the SHAP algorithm has been widely used to calculate quantitatively the contribution of each feature to a particular task[5,21]. Xiong *et al.*[5] found that critical SHAP values exist in some features when studying the hardness and ultimate tensile strength of complex concentrated alloys (CCAs) by ML. The critical feature value separates the SHAP values into positive and negative regions. This means that the feature values in the positive/negative SHAP value region improve/impair the mechanical properties of CCAs, thereby providing a straightforward assessment of the design of CCAs with high hardness and strength. Obviously, the application of SHAP, including pure and interaction SHAP values, will further promote the development of materials informatics.
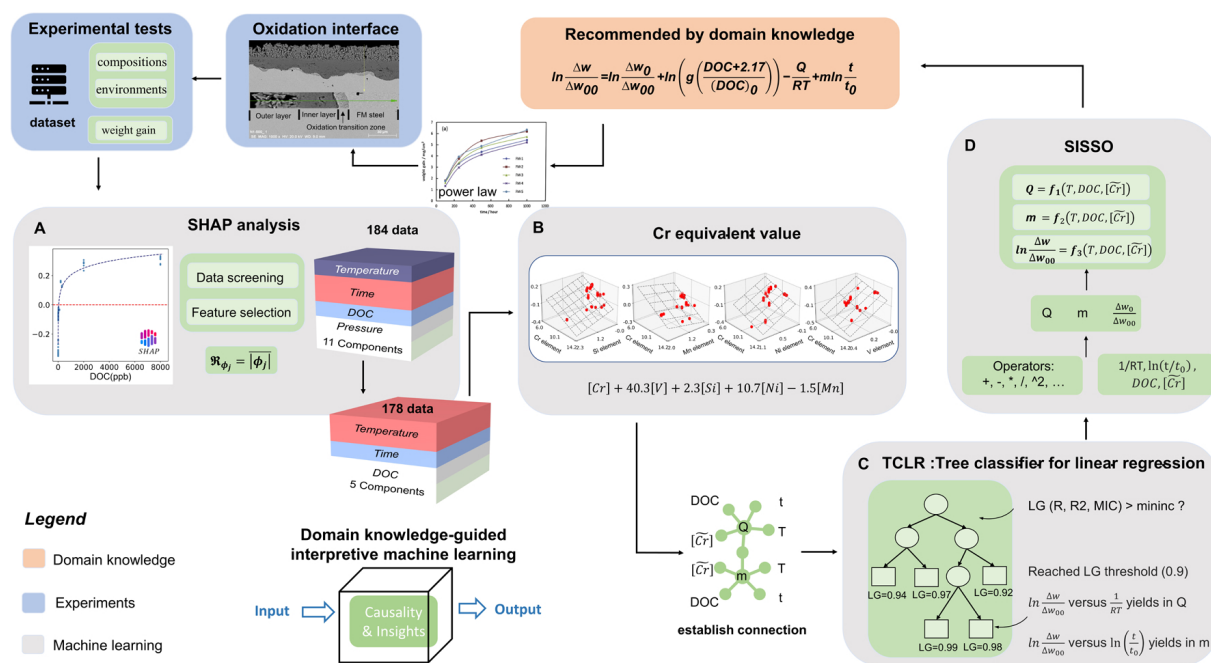
Figure 1. Domain knowledge-guided interpretive ML strategy. A: Feature selection with $\mathfrak{R}_{\phi_j}$ and data screening through 3σ criterion. B: Novel oxidation Cr equivalent of FM steels derived from joint contributions of elements. C: A prototype "divide-and-conquer" algorithm, TCLR, proposed for capturing the influences of features on time exponents and activation energy. D: Formulas of crucial terms derived by the SISSO under the constraint of prior domain knowledge.

The generalized formula established in this study accurately predicts the oxidation behavior of experimental FM steels with different alloying elements in various SCW testing conditions. Figure 1 outlines the domain knowledge-guided interpretive ML strategy, where the hub is the domain knowledge suggested Arrhenius equation. The feature importance of SHAP $\mathfrak{R}_{\phi_j} = \overline{|\phi_j|}$ is applied in feature selection effectively. Outliers are screened out based on a calculated 99.7% confidence interval [Figure 1 A] of the ML model. The oxidation Cr equivalent concentration is introduced based on the interactive decomposition of SHAP values to interpret the joint contributions of alloying elements to the oxidation behavior [Figure 1B]. The TCLR algorithm is an efficient classifier to locate samples to leaves with each sharing the same activation energy and/or time exponent [Figure 1C]. The SISSO[41] assembles the information from TCLR into a generalized formula [Figure 1D]. The Arrhenius equation is the starting point of the domain knowledge-guided interpretive ML strategy, which is a prior suggested based on our knowledge. The Arrhenius equation is also the posterior after the mining and evaluation of the ML algorithms with the experimental data, and thus transfers to a generalized formula.

METHODS

Dataset

A total of 184 oxidation data of FM steels in SCW are collected from the literature and given at the online Supplementary Information. Every datum in the FM steel oxidation (FMO) dataset includes oxidation caused weight gain in units of mg/dm^2, testing conditions, exposure time and steel chemical compositions, with a total of four testing features and 11 alloying element features (the balance element Fe is not accounted for), as shown in Table 1.

Xgboost and SHAP values

Xgboost[40] is a powerful tree-based boosting ensemble algorithm. The present work employs the Xgboost algorithm to regress the oxidation data of FM steels in SCW, and the values of hyperparameters involved are optimized by cross-validation and/or a grid search in the hyperparameter space with the open python library

Table 1. Fifteen features of FM steel oxidation data

Category	Feature name	Description
Alloying elements	Cr	Chromium (wt.%)
	Si	Silicon (wt.%)
	Mn	Manganese (wt.%)
	C	Carbon (wt.%)
	Ni	Nickel (wt.%)
	Mo	Molybdenum (wt.%)
	Nb	Niobium (wt.%)
	W	Tungsten (wt.%)
	V	Vanadium (wt.%)
	P	Phosphorus (wt.%)
	Cu	Copper (wt.%)
Testing conditions	T	Absolute temperature (K)
	Pressure	SCW pressure (MPa)
	t	Exposure time (h)
	DOC	Dissolved oxygen concentration (ppb)

scikit-learn[43]. Table S1 at Section 2.1 of the Supplementary Information lists all optimized values of the hyperparameters.

The SHAP algorithm is developed based on game theory[25]. A SHAP value, whether positive or negative, reflects the contribution of a feature to a predicted response in one datum, and the predicted response is given by an ML model. In the present work, SHAP values are calculated with Xgboost models. If there are m features of $(x_1, \cdots, x_m)$ and one predicted response $\hat{y} = f(x_1, \cdots, x_m)$, the predicted response can be expressed by the SHAP values in an additive manner as

$$\hat{y} = \phi_0 + \sum_{j=1}^{m} \phi_j, \tag{1a}$$

where ϕ_0 denotes the SHAP value of the tree root in the tree-based Xgboost algorithm and ϕ_j is the SHAP value of feature $x_j (j = 1, \cdots, m)$. The SHAP value ϕ_j can be decomposed into

$$\phi_j = \phi_{j,j} + \sum_{k \neq j, k=1}^{m} \phi_{j,k} \text{ for } (j = 1, \cdots, m), \tag{1b}$$

where the SHAP value $\phi_{j,j}$ denotes the pure contribution of feature x_j itself and $\phi_{j,k}$ is known as the interaction between features x_j and x_k, and accordingly, $\phi_{j,j}$ and $\phi_{j,k}$ are deemed the pure and interaction SHAP values, respectively. Equation (1b) indicates that the contribution from feature x_j to the predicted response is the sum of its pure contribution plus its joint contributions with other features. Therefore, a comprehensive study of SHAP values will provide insights into the mechanisms of the oxidation of FM steels in SCW.

Integration of SHAP values with domain knowledge
The game theory-based SHAP value is an additive feature attribution method, where the output is a sum of contributions of each input feature[44]. If the contributions of variables to a function are not additive in the original variable space, but additive in a mapped space, the SHAP value will be calculated in the mapped space. For the oxidation of FM steels in SCW, there are 15 features and each feature contributes to the oxidation weight gain $\Delta w = f(x_1, \cdots, x_{15})$, where $x_j (j = 1, \cdots, 15)$ represent the 15 features.

It might be inaccurate to calculate the reasonable SHAP values in the original space. Based on the domain knowledge of oxidation, we take a dimensionless Arrhenius equation of $\frac{\Delta w}{\Delta w_{00}} = \frac{\Delta w_0}{\Delta w_{00}} \exp\left(-\frac{Q}{RT}\right) \left(\frac{t}{t_0}\right)^m g\left(\frac{(DOC+r)}{(DCO)_0}\right)$. Obviously, the features T and t are explicitly shown in the Arrhenius equation. The contributions from the other features to the oxidation weight gain are via the three parameters of Δw_0, Q and m. Although the contributions of features to the oxidation weight gain are not additive in the original Arrhenius equation, the logarithmic form of $\ln\frac{\Delta w}{\Delta w_{00}} = \ln\frac{\Delta w_0}{\Delta w_{00}} + \ln\left(g\left(\frac{DOC+r}{(DOC)_0}\right)\right) - \frac{Q}{RT} + m\ln\frac{t}{t_0}$ exhibits the additive behavior and is

coherently integrated with SHAP values in the present study to explore the oxidation mechanism of FM steels in SCW. For simplicity, we let $y = \ln \frac{\Delta w}{\Delta w_{00}}$, $y_0 = \ln \frac{\Delta w_0}{\Delta w_{00}}$, $x_1 = \frac{1}{RT}$, $x_2 = \ln \frac{t}{t_0}$, $x_3 = $ DOC and other features $x_j (j = 4, \cdots, 15)$.

Feature ranking, selection and data screening
An Xgboost model is first developed with all features via ten-fold cross-validation (10-CV) and is used to calculate all SHAP values ϕ_{ij}, where subscripts "i" and "j" denote sample $i (i = 1, \cdots, n)$ and feature $x_j (j = 1, \cdots, m)$, respectively. The mean $\overline{|\phi_j|}$ of the absolute SHAP values for each feature $x_j (j = 1, \cdots, m)$ is then calculated from

$$\mathfrak{R}_{\phi_j} = \overline{|\phi_j|} = \frac{1}{n} \sum_{i=1}^{n} |\phi_{ij}|, \tag{2}$$

where $n = 184$ and $m = 15$ are the numbers of data and features, respectively. The feature importance $\mathfrak{R}_{\phi_j}$ indicates that the higher the mean of the absolute SHAP values of a feature, the greater the contribution of the feature to the response will be.

The data screening is carried out by evaluating the errors ε_i between the response y_i and the predicted response $\hat{y}_i$, viz., $\varepsilon_i = y_i - \hat{y}_i (i = 1, \cdots, n)$. The errors should follow a normal distribution $\varepsilon \sim N(0, \sigma)$. If the errors exhibit a nonzero mean, the prediction from an ML model deviates systematically from the real values. In this case, the ML model should be improved. If the errors show a normal distribution with a zero mean and few errors of data located far away from the mean, the 3σ criterion is usually taken as the threshold to gauge whether a datum is an outlier. Caution must be used in handling outliers. Repeated tests are strongly suggested to verify the reliability of the outliers and reliable outliers imply a new mechanism.

RESULTS AND DISCUSSION

Feature selection and data screening
A total of 184 data on FM steel oxidation in SCW are collected from the literature and provided in the Supplementary Information. Fifteen features are employed here and categorized into two groups, namely, alloying elements and testing conditions. The feature analysis is carried out within each of the groups. The SHAP values $\phi_{ij} (i = 1, \cdots, 184)(j = 1, \cdots, 15)$ of all features and errors $\varepsilon_i = y_i - \hat{y}_i (i = 1, \cdots, 184)$ are calculated with the Xgboost model evaluated by 10-CV. Figure 2(A)-(D) shows the SHAP values of the four testing condition features, while Figure 2(E)-(I) illustrates the SHAP values of the five alloying elements of V, Si, Cr, Ni and Mn. The SHAP values of the other six alloying elements are shown in Figure S1 in Section 2.2 of the Supplementary Information, where the blue points are calculated SHAP values and the red dashed lines indicate the SHAP value of $\phi_0 = 5.45$, which serves as a reference baseline through the origin of the vertical SHAP value axis.

As expected, Figure 2A shows that the lower the value of $\frac{1}{RT}$, the higher the SHAP value of $\phi_{\frac{1}{RT}}$, meaning that in general, the oxidation of FM steels is severer at higher temperatures. In contrast, Figure 2A also shows that at a given temperature, the SHAP value of $\phi_{\frac{1}{RT}}$ varies greatly, e.g., the SHAP value of $\phi_{\frac{1}{RT}}$ varies from 0.31 to 1.22 at $\frac{1}{RT} = 0.13029$, from 0.93 to 1.22 at $\frac{1}{RT} = 0.13775$ and from -2.04 to -1.52 at $\frac{1}{RT} = 0.17868$. The variation in $\phi_{\frac{1}{RT}}$ at a certain temperature is caused by other features, which will be analyzed later. The SHAP value of $\phi_{\ln\left(\frac{t}{t_0}\right)}$ also exhibits the anticipated trend, as illustrated in Figure 2B, where the larger the value of $\ln\left(\frac{t}{t_0}\right)$, the higher the SHAP value of $\phi_{\ln\left(\frac{t}{t_0}\right)}$. This result is logical because a longer exposure time increases the oxidation weight gain. Furthermore, the variation in $\phi_{\ln\left(\frac{t}{t_0}\right)}$ at a given value of $\ln\left(\frac{t}{t_0}\right)$ is induced by other features and will be studied later.

The SHAP value of ϕ_{DOC} behaves interestingly, as shown in Figure 2C, where at the deaerated condition (DOC
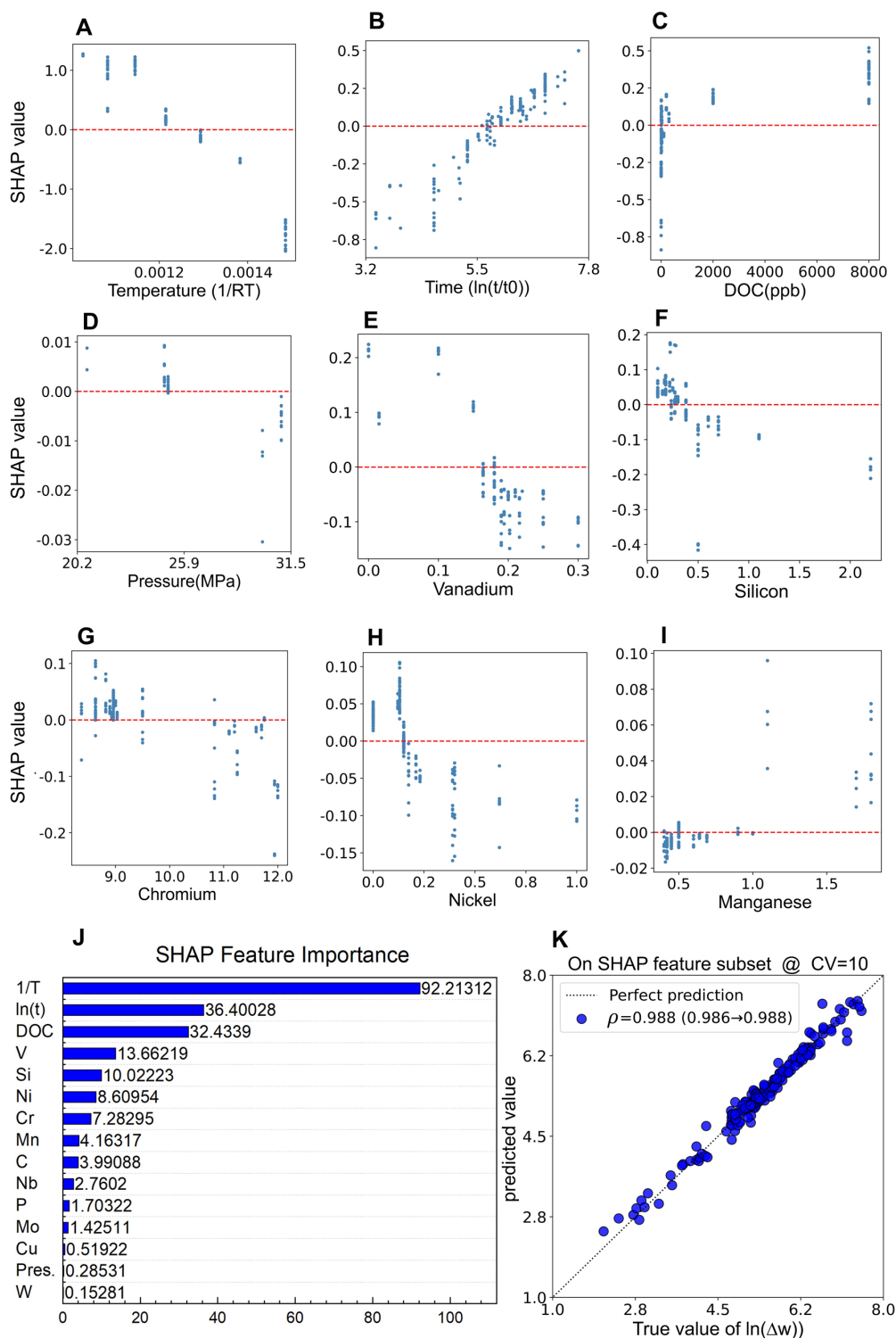
Figure 2. SHAP analysis of features. A-D: SHAP values of testing conditions, i.e., temperature, time, DOC and pressure. E-I: SHAP values of alloy compositions, i.e., V, Si, Cr, Ni and Mn. J: Feature importance ranking by $\Re_{\phi_j}$. K: Fitting effect of Xgboost model with selected eight features for 10-CV.

< 10 ppb), the SHAP value of ϕ_{DOC} changes greatly from −0.84 to 0.17. Once DOC > 10 ppb, the higher the DOC, the larger the SHAP value of ϕ_{DOC}. The shape of the SHAP values of ϕ_{DOC} suggests a logarithm function of $g\left(\frac{(DOC+r)}{(DCO)_0}\right)$. The result is rational in thermodynamics and kinetics because a higher DOC results

in a stronger driving force for oxidation. Figure 2D indicates that the magnitude of the SHAP values of ϕ_{pressure} is one or two orders smaller in comparison with $\phi_{\frac{1}{RT}}$, $\phi_{\ln\left(\frac{t}{t_0}\right)}$ and ϕ_{DOC}, ranging from −0.013 to 0.009 except for one point at −0.03 with a pressure of 30 MPa. This indicates that between the pressure range of 20.7–31 MPa, the pressure of SCW has a negligible contribution to the oxidation of FM steels in SCW.

It is somewhat surprising that vanadium plays the most important role among the studied 11 alloying elements in the oxidation of FM steels in SCW, as shown in Figure 2E, where the SHAP value of ϕ_V statistically decreases from 0.22 to −0.15 as the V content increases from 0 to 0.3 wt.%. When the V content is higher than 0.19 wt.%, the SHAP values of ϕ_V are all negative regardless of the great variation, thereby indicating that the V content should be higher than the critical value in anti-oxidation FM steels in SCW. The performance of silicon in the oxidation of FM steels in SCW is similar to that of vanadium. Figure 2F shows that the critical Si concentration is 0.38 wt.%, beyond which the SHAP values of ϕ_{Si} are all negative. However, for 0.5 wt.% Si, a few SHAP values of ϕ_{Si} appear as outliers, which will be analyzed later. Although chromium is the basic element of FM steel against oxidation, its anti-oxidation performance is unsatisfactory with a SHAP value of $\phi_{\text{Cr}} > 0$ if its content is less than 9.5 wt.%, as shown in Figure 2G. Nickel has high corrosion resistance to acids and alkalis at higher temperatures [45] and thus alloying Ni with a content higher than 0.13 wt.% into FM steels enhances the oxidation resistance of FM steels in SCW, as shown in Figure 2H. Manganese might be detrimental to the oxidation of FM steels in SCW when its content is high. Figure 2I indicates that the SHAP value of $\phi_{\text{Mn}} > 0$ if the Mn concentration is higher than 0.9 wt.%.

The feature importance defined in the SHAP method (see Methods), $\Re_{\phi_j}$, is calculated for all 15 features and plotted in Figure 2J. The three features of the testing conditions, $\frac{1}{RT}$, $\ln\left(\frac{t}{t_0}\right)$ and DOC, rank as the top three in the feature importance, indicating that the oxidation environment and time play the most important roles in the oxidation of FM steels in SCW. In contrast, the influence of SCW pressure in the range of 20.7–31 MPa on the oxidation of FM steels is trivial. In general, the feature importance of alloying elements is lower than the three features of the testing conditions, as shown in Figure 2J.

The feature selection is then conducted by the sequential backward selector wrapped with Xgboost and 10-CV, which yields the three testing features of $\frac{1}{RT}$, $\ln\left(\frac{t}{t_0}\right)$ and DOC, and five alloying elements of V, Si, Ni, Cr and Mn, i.e., the top eight in the feature importance order. It seems that the elements of V, Si and Ni play more important roles than Cr based on the feature importance ranking. As described later in the pure SHAP value analysis, Cr is the major element when its concentration is higher than 11 wt.% against FM steel oxidation in SCW. Furthermore, the errors ε_i between the response y_i and the predicted response $\hat{y}_i$, i.e., $\varepsilon_i = y_i - \hat{y}_i (i = 1, \cdots, n)$, are calculated for the original 184 data and follow a normal distribution with a zero mean and standard deviation of $\sigma = 0.35$, as shown in Figure S2 in Section 2.3 of the Supplementary Information. However, there are six errors located outside of $\pm 3\sigma$ of ε_i, meaning that the associated six data are outside the 99.7% confidence interval. These six data are removed from the following analysis and highlighted in the FMO dataset, and hereafter, only the surviving 178 data are used to explore analytic formulas for the oxidation of FM steels in SCW (the details regarding the outliers are given in Section 2.4 of the Supplementary Information). With the selected eight features, the Xgboost model predicts the values of $\ln\frac{\Delta w}{\Delta w_{00}}$ for the 178 data. The predicted values are plotted against the real values in Figure 2(K), showing an excellent prediction with a Pearson correlation value of $\rho = 0.988$, representing a slight enhancement from the original $\rho = 0.986$ with all features.

Pure SHAP values and oxidation Cr equivalent

The SHAP value of each feature is decomposed into its pure SHAP value and the interaction SHAP values, as stated in Eq. (1b). Figure 3A-H shows the pure SHAP values of the selected eight features and there are 178 pure SHAP values in each figure. A comparison of Figure 3A-H to the corresponding Figure 2A-C and Figure
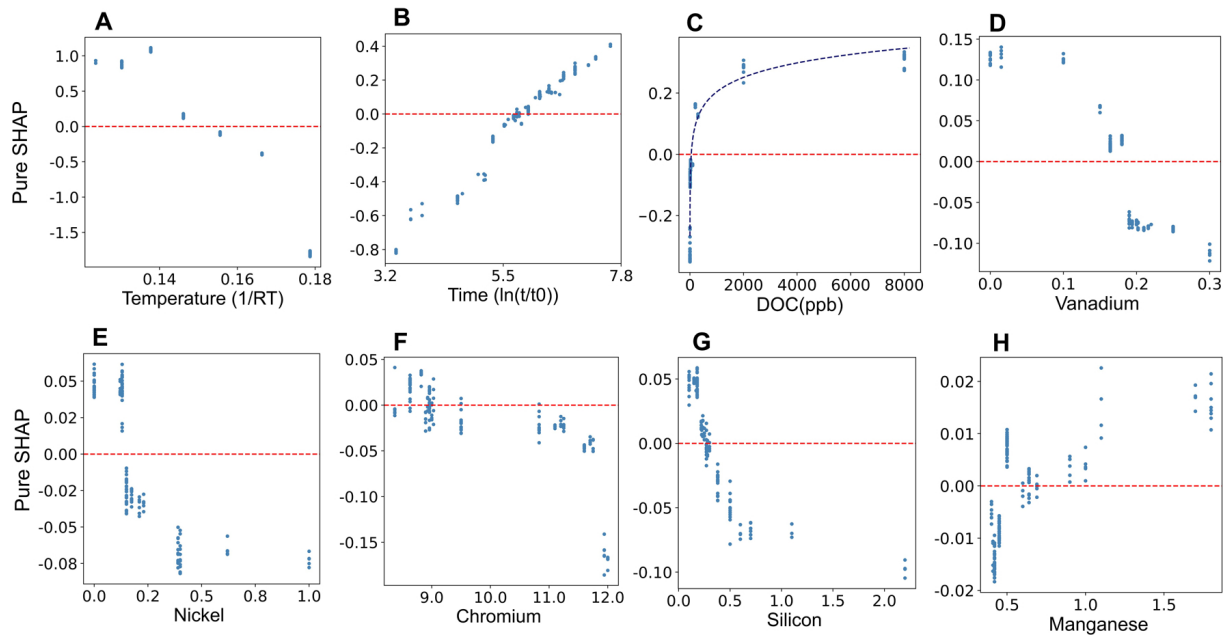
Figure 3. Pure SHAP value analysis of important features. A-C: Pure SHAP values of testing conditions, i.e., temperature, time and DOC. D-H: Pure SHAP values of alloy compositions, i.e., V, Ni, Cr, Si and Mn.

E-I indicates that for a certain feature value, the pure SHAP value scattering is much smaller than the SHAP values. This is an expected result because a pure SHAP value eliminates all interaction SHAP values from its parent SHAP value. If pure SHAP values are calculated from a perfect model of a single function of variables, the pure SHAP value of a feature will correspond to the feature value one-to-one, i.e., for one feature value, there is only one pure SHAP value.

Figure 3A shows almost ideal pure SHAP values of feature $\frac{1}{RT}$. There are [0.90, 0.93], [0.83, 0.93], [1.06, 1.11], [0.12, 0.18], [−0.12, −0.09], [−0.40, −0.38] and [−1.84, −1.76] pure SHAP values at $\frac{1}{RT}$ = 0.123598, 0.130292, 0.137753, 0.14612, 0.15557, 0.166327 and 0.178681, respectively, and the pure SHAP values at a given value of $\frac{1}{RT}$ are almost the same such that the data are condensed to one point, as shown in Figure 3A. The near perfect result implies that the oxidation of FM steels in SCW is indeed a thermally active process and the pure SHAP value of feature $\frac{1}{RT}$ catches the oxidation mechanism. Figure 3B illustrates the pure SHAP values of feature $\ln\left(\frac{t}{t_0}\right)$, where there are 36 values of $\ln\left(\frac{t}{t_0}\right)$ and the maximum of the pure SHAP values is 0.41 at $\ln\left(\frac{t}{t_0}\right)$ = 7.600902.

There are a few reasons causing multiple pure SHAP values at a given feature. The first reason might be experimental errors, which measure the degree of the experimental scattering of repeated tests. The second reason might be attributed to the Xgboost model, which approximately estimates the response from the input feature data rather than a perfect function. The third reason might be the method used to calculate the SHAP values from a tree-based algorithm (Tree-Explainer model). Figure 3C shows that the pure SHAP values of the DOC feature can be expressed by a logarithm function of $0.0682 \ln\left(\frac{(DOC+1)}{(DCO)_0}\right) - 0.267$, which is plotted by the dashed curve. The logarithm function of pure SHAP requires its argument to be larger than zero but beyond the restriction of a specific r value; hence $r = 1$ is arbitrarily assigned in the figure, and the real value of r for the oxidation in the deaerated SCW is determined in the following analysis. A similar trend is exhibited in the pure SHAP values of the element features. The outlier suspected points at 0.5 wt.% Si disappear in the pure SHAP values of the Si feature, thereby indicating they are caused by the interaction SHAP values. As expected,

the critical values of 0.19 wt.% V, 0.38 wt.% Si, 9.5 wt.% Cr, 0.13 wt.% Ni and 0.9 wt.% Mn remain unchanged in the pure SHAP values. Figure 3D shows that the minimum pure SHAP value of V is −0.122 at its highest concentration (0.3 wt.%) in the dataset, while Figure 3F indicates that the minimum pure SHAP value of Cr is −0.186 at its highest concentration of 12 wt.%. The pure SHAP values based on the available data illustrate that Cr is still the basic alloying element for anti-oxidation FM steels in SCW.

From the pure SHAP value of an individual feature, we defined the joint SHAP value of two features as

$$\phi_{jk} = \phi_{j,j} + 2\phi_{j,k} + \phi_{k,k} \text{ for } (j, k = 1, \cdots, m), \tag{3a}$$

which measures the joint contribution of two features. In general, ϕ_{jk} is a function of features x_j and x_k. As described above, the present industrial practice still employs Cr as the basic alloying element in manufacturing anti-oxidation FM steels in SCW. Therefore, Cr is taken as the basic feature x_j and V, Si, Ni and Mn are treated as the partner features x_k in the calculation of the joint SHAP value. A linear function is proposed to approximately estimate the interaction function, i.e., $\phi_{jk}(x_j, x_k) \approx \phi_{jk0} + a_j^0 x_j + a_k^0 x_k$. The two coefficients a_j^0 and a_k^0 represent the contribution weights of features x_j and x_k, respectively, to their interaction function. Then letting $a_j^0 \equiv 1$ introduces an equivalent coefficient $a_k = a_k^0/a_j^0$ for feature x_k and named it x_j equivalent coefficient of x_k. With the approach, we calculate the Cr equivalent coefficient for each of V, Si, Ni and Mn and then sum up the four Cr equivalent coefficients to yield the oxidation Cr equivalent concentration, $[\widetilde{Cr}]$,

$$[\tilde{C}r] = [Cr] + a_V[V] + a_{Si}[Si] + a_{Ni}[Ni] + a_{Mn}[Mn] \tag{3b}$$

where [·] represents the real concentration in wt.% of an element. Figure 4A-D shows the joint SHAP values of two features, ϕ_{CrSi}, ϕ_{CrMn}, ϕ_{CrNi} and ϕ_{CrV}, respectively, and each of the joint SHAP values of two features contains 178 data. Using the linear approximation and the approach mentioned above, we can then obtain the oxidation Cr equivalent concentration

$$[\tilde{C}r] = [Cr] + 40.3[V] + 2.3[Si] + 10.7[Ni] - 1.5[Mn] \tag{3c}$$

Hereafter, we use one feature of the oxidation Cr equivalent concentration to replace the five element features. Thus, the total number of features is reduced to four, one chemical composition feature and three testing condition features. With the four features, the Xgboost model is retrained with 10-CV. The predictions on the 178 data are plotted in Figure 4E, showing a perfect fitting with $\rho = 0.985$. With the Xgboost model, we calculate the pure SHAP value of the oxidation Cr equivalent concentration, $\phi_{[\tilde{C}r],[\tilde{C}r]}$ and plot it in Figure 4F. The pure SHAP value $\phi_{[\tilde{C}r],[\tilde{C}r]}$ is almost linearly proportional to the oxidation Cr equivalent concentration $[\widetilde{Cr}]$ with a correlation of $\rho = -0.97$. Figure 4F shows that when $[\widetilde{Cr}] > 19.29$ wt.%, the pure SHAP values $\phi_{[\tilde{C}r],[\tilde{C}r]}$ are all negative. It is important to consider that the oxidation behavior of FM steels in SCW depends on both the composition and microstructure of the steels. The oxidation chromium equivalent concentration is based only on the chemical composition and therefore $[\widetilde{Cr}]$ is a purely compositional representative, which will provide guidance in the design of anti-oxidation FM steels.

Activation energy and time exponents

The oxidation mechanism of FM steels in SCW is embodied in the activation energy and time exponent[29,46], which are the coefficients of $\ln \frac{\Delta w}{\Delta w_{00}}$ versus features $\frac{1}{RT}$ and $\ln\left(\frac{t}{t_0}\right)$, respectively. Figure S3 in Section 2.5 of the Supplementary Information shows the plots of $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $\frac{1}{RT}$ and $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $\ln\left(\frac{t}{t_0}\right)$, respectively, indicating multiple values of $\ln \frac{\Delta w}{\Delta w_{00}}$ for a given value of $\frac{1}{RT}$ or $\ln\left(\frac{t}{t_0}\right)$ due to the contributions of other features. To investigate the effects of other features on the activation energy and time exponent, a novel "divide-and-conquer" approach is developed based on a tree model and hence named TCLR. Maximizing the information gain in conventional tree classifiers determines the splitting feature and the associated splitting
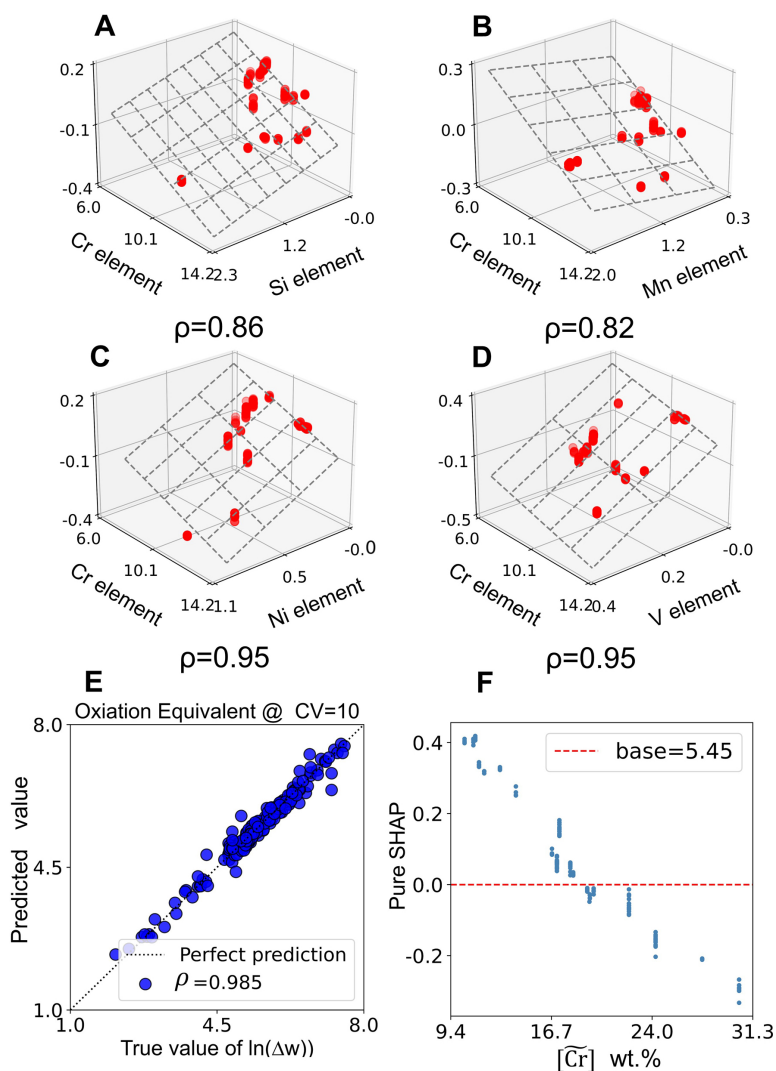
Figure 4. Joint SHAP value of two features and the derived oxidation Cr equivalent concentration. A-D: Joint SHAP value of two features, i.e., Cr with Si, Mn, Ni and V, respectively. E: Predicted values of Xgboost model versus experimental values with the transferred four features. F: Pure SHAP analysis of oxidation Cr equivalent concentration feature.

value. TCLR is developed based on that the response $y = f(x_1, \cdots, x_m)$ is a linear function of features $x_{\hat{j}}$ $(\hat{j} = 1, \cdots, \hat{m})$ $\hat{m} \leq m$, where features $x_{\hat{j}}$ $(\hat{j} = 1, \cdots, \hat{m})$ are called the linear-features. The linear-feature set is a subset of the feature set $(x_1, \cdots, x_m)$. A TCLR tree is a binary classification tree, where only the linear relationship between the response and one linear-feature is classified by using all features $(x_1, \cdots, x_m)$ in splitting. In the present work, there are two linear-features of $x_{\hat{j}=1} = \frac{1}{RT}$ and $x_{\hat{j}=2} = \ln\left(\frac{t}{t_0}\right)$ and one response $y = \ln\frac{w}{w_{00}} = f\left(\frac{1}{RT}, \ln\left(\frac{t}{t_0}\right), DOC, [\tilde{C}r]\right)$ with four features. Thus, two TCLR trees are developed and in each of TCLR trees, the absolute change in Pearson correlation coefficient (Other metrics are embedded for nonlinear functions, as demonstrated in section 1 of SI), called the Linearity Goodness (LG), between function $y = f(x_1, \cdots, x_m)$ and a linear-feature $x_{\hat{j}}$ $(\hat{j} = 1, \cdots, \hat{m})$ on nodes is taken as a generalized loss function. The split feature and split value are determined, from n data with m features $\{x_{ij}\}(i = 1, \cdots, n)(j = 1, \cdots, m)$, by maximizing the gain of average LG in the consequent nodes. For example, if the correlation coefficients in a parent node and two child nodes are denoted by ρ^p, ρ^{c1} and ρ^{c2}, respectively, the gain of average LG is defined by $\Delta|\rho| = \frac{1}{2}|\rho^{c1} + \rho^{c2}| - |\rho^p|$ and the split feature and split value x_{IJ} are determined from

$$x_{IJ} = \underset{x_{ij}}{\arg\max} \, \Delta|\rho| \, (i = 1, \cdots, n)(j = 1, \cdots, m) \tag{4}$$

The TCLR tree must be pruned, otherwise, many leaves contain only two data per leaf, which destroys the model generalization considerably. A threshold of data number might also be introduced to prune the tree, i.e., the amount of data in a leaf should not be smaller than a pre-set threshold (default minsize = 3). Another threshold for the LG is introduced to pre-prune the tree. When the LG in a node is equal to or higher than a pre-set threshold, the node will become a leaf. In general, as the value of the LG threshold increases, the model complexity of our procedure tends to improve and the tree tends to overfit. The opposite behavior occurs as the LG threshold decreases. In the bias-variance trade-off, a lower value of the LG threshold corresponds to a larger value of the number threshold of data. Thus, TCLR has the function of data screening by discarding data, which cannot be described linearly (functionally) for some of the resulting leaves in such a framework. In the present work, the LG threshold is pre-set to be 0.90 for both $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $\frac{1}{RT}$ and $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $\ln \left(\frac{t}{t_0} \right)$. The details of the TCLR algorithm are given in Supplementary Information.

The entire feature space is estimated from the regions of the four features, i.e., $[\widetilde{Cr}] \in [10.38, 30.319]$ wt.%, $T \in [673.15, 973.15]$ K, $t \in [30, 2000]$ h and DOC $\in [0, 8000]$ ppb. The TCLR separates the entire feature space into many subdomains and each subdomain corresponds to one leaf, as shown in Figure S4, Table S2 and Table S3 in Supplementary Information. For $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $1/RT$, the results indicate that 17 leaves pass the LG threshold and two fail, while for $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $\ln \left(\frac{t}{t_0} \right)$, the leaves that pass and fail are 28 and 7, respectively. The amount of data in each passed leaf ranges from 3 to 30 for $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $1/RT$ and from 3 to 10 for $\ln \frac{\Delta w}{\Delta w_{00}}$ vs. $\ln \left(\frac{t}{t_0} \right)$, indicating that reliable values of activation energy Q and time exponent *m* will be statistically obtained by TCLR. The ideal case for the measurement of activation energy Q (or time exponent *m*) is that the tests are carried out at various temperatures (or times) for a certain FM steel under other fixed testing conditions. The TCLR realizes such ideal cases from the complex data by partitioning the whole feature space of FM steels into many small rectangle domains, with one domain for a passed leaf. The samples on each passed leaf may have the same feature or varied feature values, but the variation must be too small to influence the measured activation energy Q (or time exponent m). On each of the failed leaves, the feature values of samples vary greatly so that they will not follow the same mechanism of oxidation, or it is too narrow to cover enough data subject to the capacity of the FMO dataset. One passed leaf and one failed leaf are taken as examples to illustrate the typical cases, as shown in Figure 5A-D.

Figure 5A shows one passed leaf for $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $1/RT$ and the feature values, where four data all have the same DOC = 8 ppb, three data have $[\widetilde{Cr}] = 17.27$ wt.% and one datum with $[\widetilde{Cr}] = 18.27$ wt.%, two data with $t = 40$ h and two data with $t = 50$ h. The linear fitting yields the statistically reasonable activation energy Q = 47.38 kJ/mol. Figure 5C shows one completely ideal passed leaf for the linear fitting of $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $\ln \left(\frac{t}{t_0} \right)$, yielding time exponent *m* = 0.412, where all five data have the same values of $[\widetilde{Cr}] = 12.95$ wt.%, DOC = 200 ppb and $T = 823.15$ K. There are four data on the failed leaf shown in Figure 5B, where the datum with $t = 2000$ h makes the LG not qualified. Although there are two data in Figure 5B that share the same feature values except for temperature, the two data do not allow to stay in a leaf because the threshold of data number is three. Figure 5D shows a failed leaf, where except for time, two data share the same feature values and another shares the same feature values. For the same reason that two data do not allow to stay in one leaf, the four data are restricted on a leaf with a low LG value.

The TCLR yields the values of activation energy Q and time exponent *m* as functions of the four features. The activation energy Q varies from 43.38 to 139.18 kJ/mol in the region $R_Q = \{[\widetilde{Cr}] \in [10.38, 30.319]$ wt.%, DOC $\in [0, 8000]$ ppb, $T \in [673.15, 973.15]$ K and $t \in [30, 2000]$ h}, where region R_Q is the union set
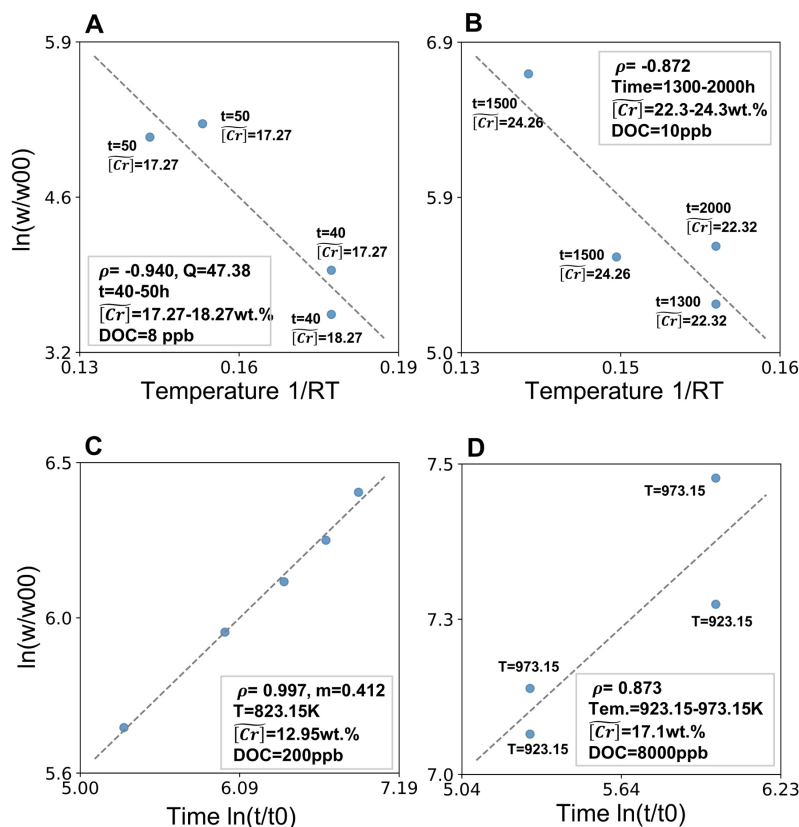
Figure 5. Data located on the leaf of TCLR. (A), (B) One passed leaf and one failed leaf on TCLR of activation energy. (C), (D) One passed leaf and one failed leaf on TCLR of time exponents.

of all leaf domains in the TCLR on $\ln \frac{\Delta w}{\Delta w_{00}}$ versus $1/RT$ withholds the inner interval it covers, containing all 178 data, including seven failed data. Similarly, the time exponent m varies from 0.112 to 0.71 in the region $R_m = \{[\widetilde{Cr}] \in [10.38, 30.319]$ wt.%, DOC $\in [0, 8000]$ ppb, $T \in [673.15, 923.15]$ K, and $t \in [30, 2000]$ h$\}$, containing 176 data including 27 failed data after removing an outlier highlighted on the red leaf of three data in Figure S4 in the Supplementary Information. The outlier gives $m = 6.49$, which is too larger than other m values. Two data on the red leaf share the same feature values except for time, which yields an unreasonable m of 15.398. This outlier should be carefully examined by experiments. The datasets contain data with four determinate features, and their corresponding activation energies and time exponents are extracted from the original FMO dataset by TCLR.

The SISSO algorithm [41], with minimization of mean absolute percentage error (MAPE, see Supplementary Information) is carried out to find analytic expressions of activation energy Q and time exponent m as functions of the four features of T, t, DOC and $[\widetilde{Cr}]$, which yields

$$Q^{\text{SISSO}}(\text{kJ/mol}) = 0.084 \frac{[\widetilde{Cr}]^2 - [\widetilde{Cr}] + \text{DOC}}{\exp(\text{DOC}/T)} + 45.09 \tag{5a}$$

$$m^{\text{SISSO}} = 0.323 - 0.061 \frac{\exp(\text{DOC}/T)}{[\widetilde{Cr}] - \sqrt{[\widetilde{Cr}] - \text{DOC}}} \tag{5b}$$

Note that $[\widetilde{Cr}]$, T and DOC are all dimensionless as pre-divided by the reference values of 1 wt.%, 1 K and 1 ppb, respectively, and the MAPE = 12.97% for Q^{SISSO} and MAPE = 39.13% for m^{SISSO}. It is interesting that feature t does not show up in both Q^{SISSO} and m^{SISSO}, meaning that time does not statistically play an important role in

activation energy Q and time exponent m. As expected, the higher the oxidation Cr equivalent concentration $[\widetilde{Cr}]$, the larger the activation energy and the smaller the time exponent. The effects of DOC and T on Q and m are not completely clear in Eqs. (5a) and (5b) and deserve more systematic investigations.

Oxidation kinetic equations

As mentioned above, the pure SHAP values of the DOC feature suggest that the oxidation kinetic in logarithmic space yield in the form of $\ln\left(\frac{\Delta w}{\Delta w_{00}}\right) = \ln\frac{\Delta w_0}{\Delta w_{00}} + \ln\left(\frac{DOC+r}{(DCO)_0}\right) - \frac{Q}{RT} + m\ln\left(\frac{t}{t_0}\right)$. The intersection of region R_Q and region R_m, $R_{Qm} = R_Q \cap R_m$, covers $R_{Qm} = [\widetilde{Cr}] \in [10.38, 30.319]$ wt.%, $T \in [673.15, 923.15]$ K, $t \in [30, 2000]$ h and DOC $\in [0, 8000]$ ppb , which contained 176 data in the FMO dataset. Two data in $[\widetilde{Cr}] = 17.0882$ wt.% are excluded in the following analysis due to their high testing temperature $T = 973.15$ K, which are out of the region R_m. With Eqs. (5a) and (5b), the values of $-\frac{Q}{RT}$ and $m\ln\left(\frac{t}{t_0}\right)$ are estimated on the whole region R_{Qm} and consequently, the values of $\ln\frac{\Delta w_0}{\Delta w_{00}} + \ln\left(\frac{DOC+r}{(DCO)_0}\right)$ are obtained on the entire region R_{Qm} by using the experimental data $\ln\left(\frac{\Delta w}{\Delta w_{00}}\right)$. Letting $\alpha = \frac{\Delta w_0}{\Delta w_{00}}\left(\frac{DOC+r}{(DCO)_0}\right)$ and plotting α versus DOC under the condition of DOC $\to 0$ yield slope $\frac{\Delta w_0}{\Delta w_{00}}$ and intersection $\frac{\Delta w_0}{\Delta w_{00}}\frac{r}{(DCO)_0}$, and hence the estimated value of $\frac{r}{(DCO)_0}$ equal to 2.17. The calculation details are given in Section 2.8 of the Supplementary Information.

To have an analytic expression of $\ln\left(\frac{\Delta w_0}{\Delta w_{00}}\right)$, the SISSO algorithm is conducted again, which yields

$$\ln\left(\frac{\Delta w_0}{\Delta w_{00}}\right)^{SISSO} = 0.084\left(\frac{[\widetilde{Cr}]^3}{T - DOC} - \sqrt{T + DOC}\right) + \frac{0.98([\widetilde{Cr}] - \frac{DOC}{T})}{\ln([\widetilde{Cr}] + DOC)} + 8.543. \quad (6c)$$

Putting all the analytic expressions together gives

$$\ln\left(\frac{\Delta w}{\Delta w_{00}}\right)^{FMO} = \ln\left(\frac{\Delta w_0}{\Delta w_{00}}\right)^{SISSO} + \ln\left(\frac{DOC + 2.17}{(DCO)_0}\right) - \frac{Q^{SISSO}}{RT} + m^{SISSO}\ln\left(\frac{t}{t_0}\right). \quad (6d)$$

The analytic formula of Eq. (5d) has strong predictive power, as shown in Figure 6, with a fitting performance of $\rho = 0.91$ and MAPE $= 6.17\%$. There are three data that cannot be accurately predicted by the formula, which fall outside of the 3σ error range, and the fitting performance will be improved to $\rho = 0.94$ and MAPE $= 5.65\%$ if these three data are excluded. The three steels share the same $[\widetilde{Cr}] = 24.255$ wt.%, DOC $= 10$ ppb, testing to 600, 1000 and 1500 h and coincidentally arrive at the upbound of testing temperature (923.15 K) on domain R_{Qm}. This may lead to the slightly underfitting of Eq. (5d) on them. In addition, the graphical user interfaces of the Xgboost prediction model were developed and are available at the end of this manuscript.

CONCLUDING REMARKS

In this study, we develop a domain knowledge-guided interpretive ML strategy and demonstrate it by the discovery of the generalized formula for FM steel oxidation in SCW. The domain knowledge suggests the generalized Arrhenius oxidation formula of $\frac{\Delta w}{\Delta w_{00}} = \frac{\Delta w_0}{\Delta w_{00}}\exp\left(-\frac{Q}{RT}\right)\left(\frac{t}{t_0}\right)^m g\left(\frac{(DOC+r)}{(DCO)_0}\right)$, based on the fact that the oxidation is a thermally active process and the collected data are conducted at various temperatures. The generalized Arrhenius formula also includes a power law of exposure time to generalize all potential time dependences in the complicated oxidation process including diffusion and phase transformation.

The oxidation chromium equivalent concentration $[\tilde{Cr}] = [Cr] + 40.3[V] + 2.3[Si] + 10.7[Ni] - 1.5[Mn]$ is, for the first time, developed to quantitatively represent the contributions of alloying elements of Cr, V, Si, Ni and Mn to the oxidation of FM steels in SCW.
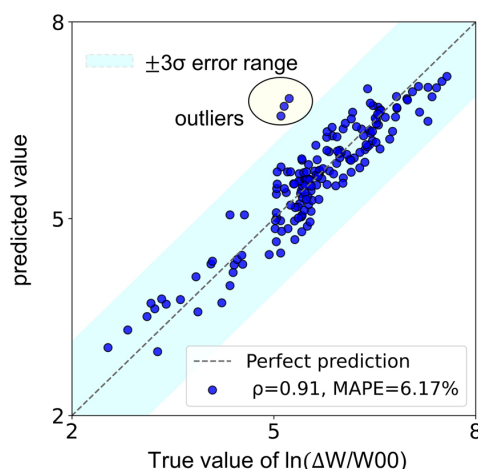
Figure 6. Predicted values of Eq. 5(d) versus experimentally measured values. Each dot represents an FM sample, and the dots on the dotted line indicate the equation predicted values are consistent with experimental observations. The light blue region covers 3σ error range, where Eq. (5d) fitted extremely well on.

The developed TCLR algorithm is scientifically significant to materials informatics because it captures linear relationships between tasks and features. It is expected that when an original feature space is mapped to a high-dimensional space, the TCLR algorithm is able to capture linear relationships in the high-dimensional space. More affords are needed to further develop the TCLR algorithm.

The generalized Arrhenius oxidation formula has very high prediction accuracy with a Pearson correlation coefficient ρ of 0.91 and a MAPE of 6.17% validated by the 176 experimental data. It is expected that more experiments will be conducted and verify the generalized Arrhenius oxidation formula for oxidation of FM steels in SCW.

DECLARATIONS

Authors' contributions
Performed the research, analysed data, wrote the programmers, and drafted the manuscript: Cao B
Collected the data. Tong-Yi Zhang and Ziqiang Dong supervised the project: Yang S, Sun A
Studied and revised the manuscript more on the oxidation mechanism: Dong Z
Designed the study, performed the research, analysed data, revised, and finalized the manuscript: Zhang TY
Discussed the results: Cao B, Yang S, Sun A, Dong Z, Zhang TY

Availability of data and materials
All experimental data collected in the study are contained in Supplementary Information (FM steel oxidation dataset) and are also available at: https://github.com/Bin-Cao/TCLRmodel. The ML methodology described in the present work was implemented in Python. Source codes of the programmers and algorithms are available at: https://github.com/Bin-Cao/TCLRmodel.

Conflicts of interest
We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Copyright
© The Author(s) 2022.

REFERENCES

1. Wei QH, Xiong J, Sun S, Zhang T-Y. Multi-objective machine learning of four mechanical properties of steels. *Sci Sin -Tech* 2021;51: 722–36. DOI
2. Xiong J, Zhang T-Y, Shi S. Machine learning of mechanical properties of steels. *Sci China Technol Sci* 2020;63:1247–55. DOI
3. Leitherer A, Ziletti A, Ghiringhelli LM. Robust recognition and exploratory analysis of crystal structures via Bayesian deep learning. *Nat Commun* 2021;12: 6234. DOI
4. Sun S, Ouyang R, Zhang B, Zhang T-Y. Data-driven discovery of formulas by symbolic regression. *MRS Bull* 2019;44:559–64. DOI
5. Xiong J, Shi S, Zhang T-Y. Machine learning of phases and mechanical properties in complex concentrated alloys. *Journal of Materials Science & Technology* 2021;87:133–42. DOI
6. Xie SR, Quan Y, Hire AC, et al. Machine learning of superconducting critical temperature from Eliashberg theory. *npj Comput Mater* 2022;8. DOI
7. Levämäki H, Tasnádi F, Sangiovanni DG, Johnson LJS, Armiento R, Abrikosov IA. Predicting elastic properties of hard-coating alloys using ab-initio and machine learning methods. *npj Comput Mater* 2022;8. DOI
8. Roy Chowdhury P, Ruan X. Unexpected thermal conductivity enhancement in aperiodic superlattices discovered using active machine learning. *npj Comput Mater* 2022;8. DOI
9. Zhu YQ, Xu T, Wei Q, et al. Linear-superelastic Ti-Nb nanocomposite alloys with ultralow modulus via high-throughput phase-field design and machine learning. *npj Comput Mater* 2021;7. DOI
10. Wang JH, Jia J, Sun S, Zhang T-Y. Statistical learning of small data with domain knowledge-sample size-and pre-notch length- dependent strength of concrete. *Engineering Fracture Mechanics* 2022;259:108160. DOI
11. Xue D, Balachandran PV, Hogden J, Theiler J, Xue D, Lookman T. Accelerated search for materials with targeted properties by adaptive design. *Nat Commun* 2016;7:11241.DOI
12. Fung V, Hu G, Ganesh P, Sumpter BG. Machine learned features from density of states for accurate adsorption energy prediction. *Nat Commun* 2021;12:88. DOI
13. Attia PM, Grover A, Jin N, et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* 2020; 578:397–402. DOI
14. Saito Y, Shin K, Terayama K, et al. Deep-learning-based quality filtering of mechanically exfoliated 2D crystals. *npj Comput Mater* 2019; 5. DOI
15. Li X, Zhao J, Cong J, et al. Machine learning guided automatic recognition of crystal boundaries in bainitic/martensitic alloy and relationship between boundary types and ductile-to-brittle transition behavior. *Journal of Materials Science & Technology* 2021;84:49–58. DOI
16. Dai F, Wen B, Sun Y, Xiang H, Zhou Y. Theoretical prediction on thermal and mechanical properties of high entropy (Zr0.2Hf0.2Ti0.2Nb0.2Ta0.2)C by deep learning potential. *Journal of Materials Science & Technology* 2020;43:168–74. DOI
17. Ramakrishna S, Zhang T-Y, Lu W, et al. Materials informatics. *J Intell Manuf* 2019;30:2307–26. DOI
18. Lookman T, Balachandran PV, Xue D, Yuan R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput Mater* 2019;5. DOI
19. Wen C, Zhang Y, Wang C, et al. Machine learning assisted design of high entropy alloys with desired property. *Acta Materialia* 2019;170: 109–17. DOI
20. Balachandran PV, Kowalski B, Sehirlioglu A, Lookman T. Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning. *Nat Commun* 2018;9:1668. DOI
21. Yan L, Diao Y, Lang Z, Gao K. Corrosion rate prediction and influencing factors evaluation of low-alloy steels in marine atmosphere using machine learning approach. *Sci Technol Adv Mater* 2020;21:359–70. DOI
22. Jablonka KM, Jothiappan GM, Wang S, Smit B, Yoo B. Bias free multiobjective active learning for materials design and discovery. *Nat Commun* 2021;12:2312. DOI

23. Garrido Torres JA, Gharakhanyan V, Artrith N, Eegholm TH, Urban A. Augmenting zero-Kelvin quantum mechanics with machine learning for the prediction of chemical reactions at high temperatures. *Nat Commun* 2021;12:7012. DOI

24. Lu L, Meng X, Mao Z, Karniadakis GE. DeepXDE: A deep learning library for solving differential equations. *SIAM Rev* 2021;63:208–28. DOI

25. Lundberg SM, Lee SI. A unified approach to interpreting model predictions, Proceedings of the 31st international conference on neural information processing systems, 2017, pp. 4768–4777. Available from: https://proceedings.neurips.cc/paper/2017/hash/8a20a86219786 32d76c43dfd28b67767-Abstract.html [Last accessed on 20 Apr 2022] DOI

26. Schulenberg T, Leung LK, Oka Y. Review of R&D for supercritical water cooled reactors. *Progress in Nuclear Energy* 2014;77:282–99. DOI

27. Zhong X, Wu X, Han E. Effects of exposure temperature and time on corrosion behavior of a ferritic–martensitic steel P92 in aerated supercritical water. *Corrosion Science* 2015;90:511–21. DOI

28. Klueh R, Nelson A. Ferritic/martensitic steels for next-generation reactors. *Journal of Nuclear Materials* 2007; 371: 37–52. DOI

29. Ampornrat P, Was GS. Oxidation of ferritic-martensitic alloys T91, HCM12A and HT-9 in supercritical water. *Journal of Nuclear Materials* 2007;371:1–17. DOI

30. Li Y, Xu T, Wang S, et al. Modelling and Analysis of the Corrosion Characteristics of Ferritic-Martensitic Steels in Supercritical Water. *Materials (Basel)* 2019; 12: 409. DOI

31. Tan L, Ren X, Allen T. Corrosion behavior of 9–12% Cr ferritic-martensitic steels in supercritical water. *Corrosion Science* 2010;52:1520–8. DOI

32. Li H, Cao Q, Zhu Z. High temperature oxidation behavior of ferritic steel in supercritical water at 550–700 °C. *Materials at High Temperatures* 2018;36:111–6. DOI

33. Zhu Z, Xu H, Jiang D, Mao X, Zhang N. Influence of temperature on the oxidation behaviour of a ferritic-martensitic steel in supercritical water. *Corrosion Science* 2016;113:172–9. DOI

34. Li Y, Wang S, Sun P, et al. Investigation on early formation and evolution of oxide scales on ferritic-martensitic steels in supercritical water. *Corrosion Science* 2018;135:136–46. DOI

35. Liu Z, Corrosion behavior of designed ferritic-martensitic steels in supercritical water. Canada: ProQuest Dissertations Publishing. 2013. DOI

36. Dong Z, Li M, Behnamian Y, et al. Effects of Si, Mn on the corrosion behavior of ferritic–martensitic steels in supercritical water (SCW) environments. *Corrosion Science* 2020;166:108432. DOI

37. Sun L, Yan W. Estimation of oxidation kinetics and oxide scale void position of ferritic-martensitic steels in supercritical water. *Advances in Materials Science and Engineering* 2017;2017:1–12. DOI

38. Bischoff J, Motta AT. Oxidation behavior of ferritic-martensitic and ODS steels in supercritical water. *Journal of Nuclear Materials* 2012; 424:261–76. DOI

39. Zhang N, Xu H, Li B, Bai Y, Liu D. Influence of the dissolved oxygen content on corrosion of the ferritic-martensitic steel P92 in supercritical water. *Corrosion Science* 2012;56:123–8. DOI

40. Chen T, Guestrin C. Xgboost: A scalable tree boosting system; proceedings of the Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, F, 2016. DOI

41. Ouyang R, Curtarolo S, Ahmetcik E, Scheffler M, Ghiringhelli LM. SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. *Phys Rev Materials* 2018; 2. DOI

42. Zhang T-Y, Cao B, Zhang S-Y, Sun S. Tree-classifier for linear regression software [No.2021SR1951267], 2021. Available from: https://register.ccopyright.com.cn/publicInquiry.html?type=softList%C2%AEisterNumber=2021SR1951267%5C&keyWord=%E4%B8%8A%E6%B5%B7%E5%A4%A7%E5%AD%A6%5C&publicityType=ALL%C2%AEisterDateType=ALL [Last accessed on 20 Apr 2022] DOI

43. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python, the journal of machine learning research 12 (2012) 2825–2830. Available from: https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?ref=https://githubhelp.com [Last accessed on 20 Apr 2022] DOI

44. Lundberg SM, Erion GG, Lee SI. Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv* 2018; 1802.03888. Doi: 10.48550/arXiv.1802.03888. DOI

45. Uusitalo M, Vuoristo P, Mäntylä T. High temperature corrosion of coatings and boiler steels below chlorine-containing salt deposits. *Corrosion Science* 2004;46:1311–31. DOI

46. Truhlar DG. Interpretation of the activation energy. *J Chem Educ* 1978;55:309. DOI