



# Prediction of Creep Life Using an Explainable Artificial Intelligence Technique and Alloy Design Based on the Genetic Algorithm in Creep-Strength-Enhanced Ferritic 9% Cr Steel

Byeong Ook Kong<sup>1</sup> · Min Soo Kim<sup>1</sup> · Byung Hoon Kim<sup>1</sup> · Je Hyun Lee<sup>1,2</sup>

Received: 24 June 2022 / Accepted: 10 September 2022 / Published online: 11 October 2022  
© The Author(s) under exclusive licence to The Korean Institute of Metals and Materials 2022

## Abstract

Creep-strength-enhanced ferritic steels have been introduced for high-temperature components. The high accuracy of creep life prediction and new high creep-strength alloys are needed for safe long-term operation. In this paper, a data-driven method was used to overcome the limitations of the simple-parameter life prediction method that does not consider the complex interactions that occur between many input variables. The explainable artificial intelligence technology, such as the Shaply additive explanation value (SHAP), enables intuitive understanding of the effect of individual variables on creep life and helps to quantitatively evaluate the effect. The artificial intelligence model was optimized using a genetic algorithm, and a method for proposing an optimal alloy composition with a desired creep life could be presented. In this paper, the maximum expected creep life of 58,640 h in the standard component range of ASME SA213 T92 under the conditions of 650 °C and 100 MPa was predicted, and a new alloy composition having a creep life of over 100,000 h under the same conditions is proposed.

**Keywords** Creep · 9% Cr steel · Explainable artificial intelligence technology · Genetic algorithm

## 1 Introduction

The 9%–12% Cr series heat-resistant steel is commonly used at high-temperature–high-pressure components of thermal power and chemical plants, such as boiler tubes, pipe, steam turbine rotor and reactor vessel [1]. These components are designed considering various types of deterioration damage mechanisms for prolonged safe operation at high temperature. Among the high-temperature damage mechanisms,

creep is one of the most basic characteristics for determining the allowable stress in the design. There is a concerted research effort for developing an alloy design with improved creep strength [2–4] and a reliable creep life prediction method [5, 6].

The Larson–Miller parameter (LMP) method is representative of various methods proposed for creep life prediction and is widely used by researchers and engineers because it can predict creep life through the simple formula [LMP = Operating Temp. × (C + log (Life))], where C is a constant [7]. However, Masuyama et al. [8] applied the LMP method to Grade 91 steel in various manufacturing processes and alloy compositions, and confirmed that the prediction accuracy strongly depends on the LMP materials' constant value. In addition, Masuyama et al. [8] proposes using C = 31 as the optimal LMP constant value.

Even when C = 31 was applied to the LMP constant value in a Grade 92 alloy, which is a similar grade produced under various composition and heat-treatment conditions, it was confirmed that a large value of the LMP prediction spread (Fig. 1). This means that the lifetime prediction method using a few simple parameters is reliable when the data used were produced within a certain range of conditions, but in the case of data obtained from various compositions and

✉ Je Hyun Lee  
ljh@changwon.ac.kr  
Byeong Ook Kong  
superkongja@changwon.ac.kr  
Min Soo Kim  
minsoo.kim@changwon.ac.kr  
Byung Hoon Kim  
bhk5830dh@changwon.ac.kr

<sup>1</sup> Research Center for Technological Innovation on Future Energy, Changwon National University, Changwon 51140, Republic of Korea

<sup>2</sup> Department of Materials Science and Engineering, Changwon National University, 51140 Changwon, Republic of Korea

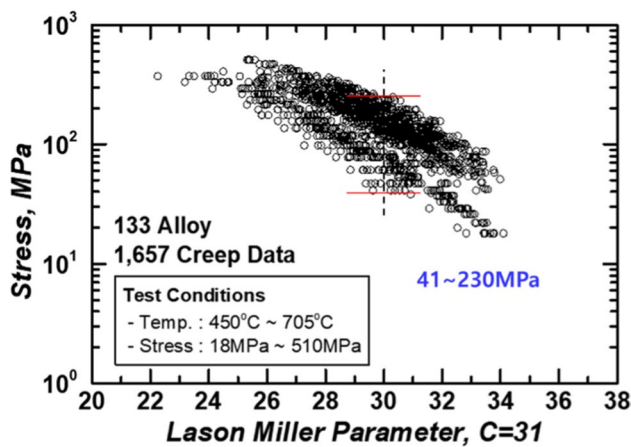


Fig. 1 Larson-Miller plot of creep rupture data

manufacturing conditions, it is judged that deterioration in prediction accuracy is unavoidable. It is difficult to consider many variables using the parameter method. In addition, because it is almost impossible to examine the interdependence between all the variables, few models can be used in complex prediction fields, such as alloy design [9].

As one of the efforts to overcome these limitations, data-based artificial intelligence (AI) learning techniques are attracting attention. Badeshia et al. [10] showed the applicability to various fields such as welding as well as creep life prediction using a neural network model. Recently, results using various machine learning models have also been reported [11, 12]. However, in the case of a data-driven model, it is difficult to understand the effect of changes in individual input features on the target like a “black-box” intuitively; hence, there is a limit to interpreting the decision-making results of the model [13].

Recently, efforts are being made to solve this problem by using explainable AI (XAI) technologies [14], which largely use either model-specific or model-agnostic techniques. Model-specific XAI methods work by inspecting or having access to the model internals. Model-agnostic methods work by investigating the relationship between input–output pairs of trained models. It is depend on the relationship between input–output pairs of trained models and are very useful for cases when there is no theory or other mechanism to interpret what is happening inside the model [15]. For example, it could provide a how an input image is reconstructed in a specific layer through layer-level visualization in convolutional nearest-neighbor (CNN) algorithm[16–18]. Partial dependence plots (PDP) and individual conditional expectation (ICE) plots give information of visualization the interaction between a target response and input features of interest [19, 20].

Local interpretable model-agnostic explanations (LIME) [21] and Shaply additive explanations (SHAP) [22] are

representative models widely used in practice. Among them, the SHAP technology [23] is a representative AI technique that can be explained by analyzing the relationship between an input variable and the model-predicted value by using the SHAP value of the individual input variable. The SHAP value used in this paper is a concept originating from a method to properly evaluate the contribution of participants to the prize money obtained through cooperation between game participants during the game. SHAP informs the importance of each input feature on the target value and also confirms information about the direction and magnitude of an effect, making it possible to easily understand and explain the machine learning model.

Model analysis was performed using the permutation feature importance tool, which is a model-agnostic technique. The permutation feature importance is defined as the decrease in a model score when a single feature value is shuffled randomly [24]. This procedure breaks the relationship between the feature and the target, and thus a drop in the model score is indicative of how much the model depends on the feature. Permutation importance does not reflect on the intrinsic predictive value of a feature by itself, but on how important this feature is for a particular model [25].

The ensemble method, which is known to exhibit excellent performance in model building and optimization, was used for the final model selection through learning [26–28]. In particular, supervised learning algorithms perform the task of searching the hypothesis space for optimization tasks to solve specific problems [29]. However, it is very difficult to find a hypothetical hypothesis in a complex multidimensional space even if the hypothesis space contains parameters that are very suitable for solving a specific problem. One useful method for solving this is ensembles, which combine multiple hypotheses to form a better hypothesis.

In this paper, the effect of alloying elements on the creep life of the ASME SA213 T92 material [30] is analyzed using XAI technology, and the maximum creep life is predicted using model optimization and genetic algorithm (GA) methods. By using an AI-based alloy design technology, based on the composition of ASME SA213 T92, we propose a new alloy composition with a creep strength of 100,000 h at 650 °C and 100 MPa. This method is to guide the next phase of experiments by gaining insights from previously completed experiments to effectively reduce the time and cost of materials discovery.

## 2 Data and Method

### 2.1 Datasets and Data Preprocess

Table 1 shows the alloy composition, heat-treatment conditions, and creep life of 133 types of martensitic heat-resistant

**Table 1** Detail of the input features

Alloy elements, wt%	Range	Mean	Alloy elements, wt%	Range	Mean
C	0.004~0.23	0.109	Si	0.01~0.86	0.300
Mn	0.27~0.92	0.506	P	0.001~0.029	0.014
S	0.001~0.02	0.008	Cr	7.915~12.9	8.776
Mo	0.04~2.99	0.905	W	0.01~3.93	0.509
Ni	0.01~2.00	0.278	Cu	0.01~0.87	0.074
V	0.01~0.28	0.146	Nb	0.005~0.312	0.043
N	0.001~0.165	0.0367	Al	0.001~0.057	0.0122
B	0.0003~0.051	0.0012	Co	0.008~2.5	0.1025
Normalizing temperature, °K	1123~1453	1012	Tempering temperature, °K	892~1133	722.296
holding time, hr	0.17~15	1.89	holding time, hr	0.5~30	
<i>Test conditions</i>					
Stress, MPa	18~510	161	Temperature, °K	723~977	603

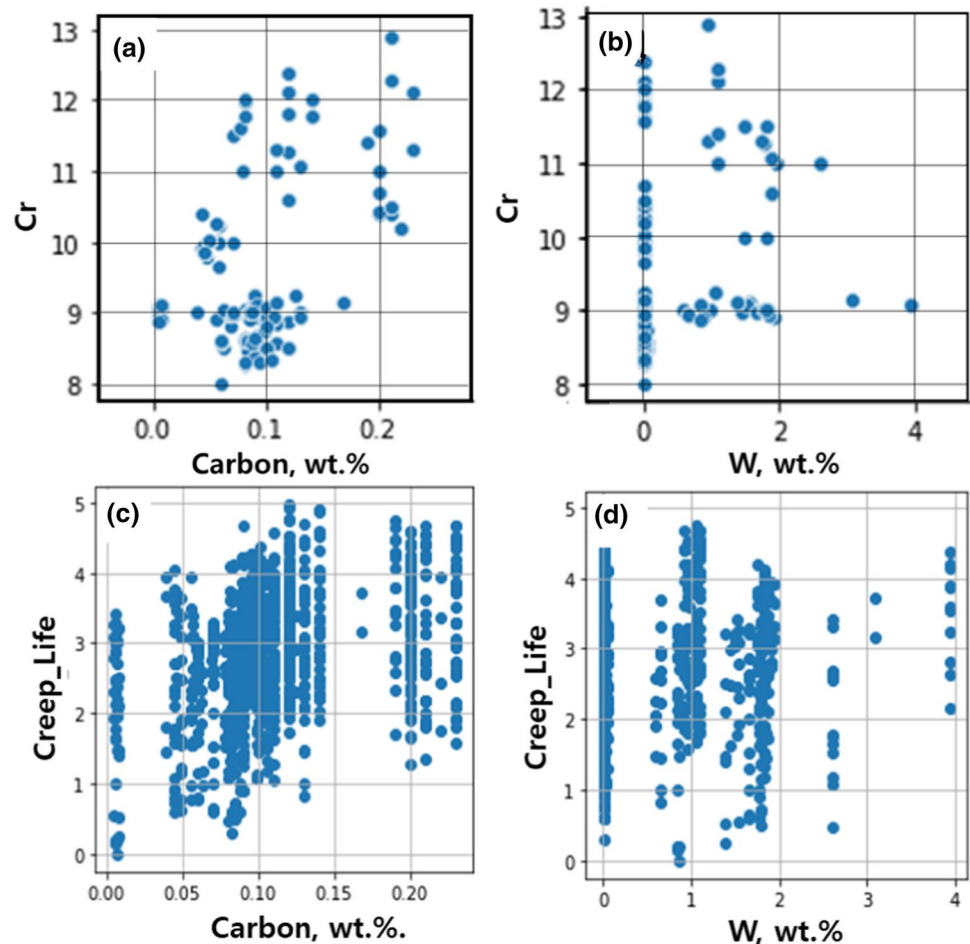
Creep Life, hrs : 0.6 ~ 167723 (average : 730)

steels used in this study with a content range of 7.9%–12.9% Cr [31]. The data consisted of 21 input features (variable) and one target (output) variable, and all data were normalized by the Min–Max scaling method to minimize the influence of the size and distribution of data values. In addition, the temperature data was used by converting it to  $T = 1000$  K considering the diffusion theory and having a relatively

larger value than the alloy component. This transformation contributes to model performance improvement [32].

To secure prediction accuracy in data-based learning, it is important to have low multicollinearity between input features. This was reviewed in the preprocessing process of the learning data and the representative results are shown in Fig. 2a and b. There is a low correlation between Cr, C,

**Fig. 2** Distributions between input features. Scatter plot between Cr and Carbon **a**, Cr and W **b**. Pair plot of Creep life with carbon and W **c**, **d**



and W alloy components, and all other features show similar trends, confirming that the issue related to multicollinearity is not significant. The distribution of creep life according to C and W content is shown in Fig. 2c, d, and the creep life is very irregularly distributed, which means that it is difficult to find a specific correlation between individual features and creep life. For actual creep life prediction, all the features must be considered at the same time, and this is almost impossible through the conventional simple regression method. The LMP stress relationship for all the data is shown in Fig. 3. The creep life distribution is large even at the same stress, which shows that in the case of the current data set, the LMP stress method has the potential to cause a large error in creep life prediction. It may be difficult to use the LMP method as a creep life prediction model because it is predicted that the creep strength can have a very wide distribution from 41 to 232 MPa under the condition that the LMP value is 30. To overcome these limitations, this work tried to interpret data using various data-driven AI methods.

## 2.2 Model Selection and Evaluation

For the optimal prediction model selection, traditional machine learning models such as the support vector machine (SVM) algorithm, the K-nearest-neighbor (KNN) algorithm, and the elastic net model were evaluated, and the random forest regressor and gradient boosting (GB) regressor were reviewed as boosting models. In addition, various latest machine learning models were used, such as the extra GB (XGB) regressor and the light GB (LGB) regressor, have shown many achievements in the data science competition field. The deep nearest-neighbor (DNN) regression model, which is a deep learning model, was also used for evaluation.

In the training process, the entire data were divided into training and test data at a ratio of 7:3, and the model was optimized by automatically changing the parameters during the

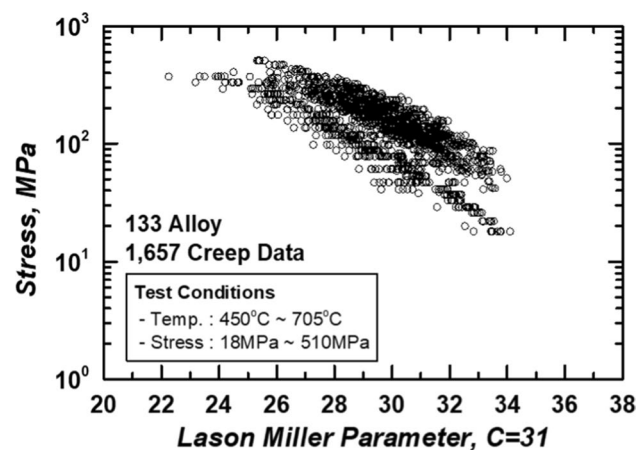


Fig. 3 Lasson-Miller plot of the data used to train the model

learning process with the pipeline hyperparameter turning method (ElasticNet model, 97,500 fit repeat; RandomForest model, 1200 fit repeat; GB model, 240 fit repeat; XGB model, 576 fit repeat; LGB model, 1200 fit repeat; DNN model: 500 fit repeat). For the DNN model, the model with three hidden layers (training  $R^2$  score, 0.922; test  $R^2$  score, 0.900) outperformed the complex model with four or more hidden layers (training  $R^2$  score, 0.922; test  $R^2$  score, 0.900) was evaluated. In the case of the DNN model, the performance strongly depends on the number of hidden layers. As a result of the evaluation, it was confirmed that the performance of the three hidden layer models was superior to that of a complex model with four or more. Thus, the data used in this paper were not sufficient for using the DNN model.

## 2.3 Ensemble Methods

The core of the ensemble model is to make a strong classifier by combining several weak classifiers, but in this work, the ensemble model was implemented using a well-trained model. The SVM (test score,  $-0.001$ ), KNN (test score,  $-0.068$ ), ElasticNet (test score,  $0.519$ ) and RandomForest (test score,  $0.630$ ) models that showed low performance during the model building process were excluded. The final ensemble model was built by selecting four models, GB, XGB, LGB, and DNN, which were evaluated for excellent performance.

## 2.4 Genetic Algorithm(GA)

In general, GAs are commonly used to generate high-quality solutions for optimizing search problems by relying on biologically inspired operators, such as mutation, crossover, and selection [33]. A GA is a metaheuristic algorithm inspired by the process of natural selection that belongs to the larger class of evolutionary algorithms. The evolution usually starts from a population of randomly generated individuals and is an iterative process, with the population in each iteration called a generation. In each generation, the fitness of every individual in the population is evaluated, where the fitness is usually the value of the objective function in the optimization problem being solved. The more fit individuals are stochastically selected from the current population and each individual's genome is recombined and possibly randomly mutated to form a new generation. The new generation of candidate solutions is then used in the next iteration of the algorithm [34].

**Table 2** The main parameters of the selected artificial intelligence model and the  $R^2$  values of the train and test data prediction

Model	Parameters	$R^2$ value	
		Train	Test
DNN (D)	Hidden layer: 26, 29, 31 Max.Iter. = 2,000,000, learning	0.922	0.900
XGB (X)	Max pepth: 2, No. of estimators = 5000	0.999	0.943
GB (G)	Max depth = 2 No. of estimators = 5000	0.992	0.936
LGB (L)	Max depth = 2 No. of estimators = 2800	0.924	0.88

**Table 3** Ensemble model configuration and  $R^2$  value of each model

Model	Ensemble	$R^2$ value	
		Train	Test
Ensemble 1	DNN	0.922	0.900
Ensemble 2	DNN + XGB	0.979	0.943
Ensemble 3	DNN + XGB + GB	0.987	0.947
Ensemble 4	DNN + XGB + GB + LGB	0.982	0.942

### 3 Results and Analysis

#### 3.1 Model Building and Parameter Turning

After learning and parameter fine-tuning using various models, the model accuracy was evaluated with the

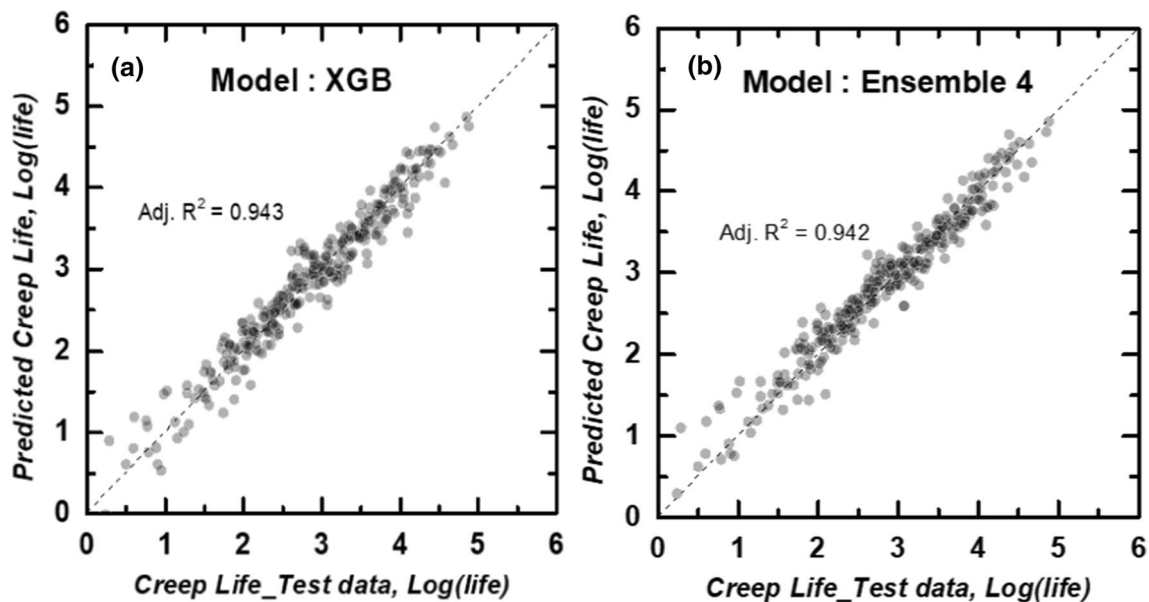
coefficient of determination and the  $R^2$  value. The top four selected models are shown in Table 2. Because the training  $R^2$  scores for all models were over 0.9 and the test scores were over 0.88, the selected model was considered to be well-trained. In particular, the XGB model, which is a representative tree-based model, was confirmed to have a very high training  $R^2$  value of 0.99; hence, there was a risk of overfitting. However, the predicted test  $R^2$  value was also fairly high at 0.943; thus, overfitting did not occur seriously. Nevertheless, it was necessary to lower the  $R^2$  value.

As mentioned earlier, the final ensemble model was built using four single models (Ensemble 4) to reduce the overfitting tendency of the single model and to improve the model generality. Model optimization was performed through the ensemble method so that the excessively high training  $R^2$  value of a single model was lowered without a decrease in the test  $R^2$  value. A model evaluation index and results are shown in Table 3. The Ensemble 4 model improved both the training and test  $R^2$  scores based on the DNN model. Compared with the XGB model, it was possible to slightly reduce the training  $R^2$  score without decreasing the test  $R^2$  score.

The results of predicting the creep life of the test data using the single XGB model and the Ensemble 4 model are shown in Fig. 4. Both models predict the creep life well except for the short-term creep range of about 100 h.

#### 3.2 Model Analysis Using Explainable Artificial Intelligence

To analyze the effect of individual features on the target value, the Pearson correlation coefficient (PCC) [35]

**Fig. 4** Predicted creep life using a final fine-tuned model. **a** XGB model, **b** Ensemble 4 model

considering the degree of linear correlation between two variables and the Spearman correlation coefficient that does not require the assumption of a linear relationship between the two variables has been used widely [36]. However, these analysis methods are useful in low-dimensional data space, but there are limitations in multidimensional analysis. Hence, efforts are being made to solve these limitations through dimension-reduction techniques, such as principal component analysis (PCA).

In this work, the effect of input features on creep life feature importance was evaluated using XAI technology and it was intended to be used in the alloy development process as well as in model understanding and life prediction.

Feature importance was evaluated using SHAP and permutation values based on the XGB and Ensemble 4 models, and the results are shown in Fig. 5a and b. The x-axis of Fig. 5a represents SHAP value magnitudes over all samples, and the base value means the average creep life of the training data set. The y-axis shows the mean absolute SHAP values ranging from top to bottom for the entire dataset, i.e., the average magnitude of each variable’s impact on the predicted creep life for all instances. How much impact on the model’s prediction for this the distribution of the impacts each feature has on the model output. As the input feature value increases, it is expressed in red (blue) when the SHAP value increases (decreases). For example, in the case of the Test\_Stress feature with blue color, if the test

stress is increased, the creep life can be reduced by up to 3.5 SHAP value ( $10^{3.5}$ ) or more. Conversely, if it is decreased, the creep life is increased by more than 4 SHAP value ( $10^4$ ). In particular, increasing the normalization temperature (the N\_Temp value is decreased in the figure) can significantly increase the creep life.

As a result of the feature importance analysis, in the case of the data set used in this work, it was evaluated that the normalizing temperature had the strongest effect on creep life, and that the alloying elements affected the creep life in the order of V, W, Mo, C, and Si. It was thought that it could be used as a very useful technology for the understanding of and developing alloy properties in the future.

### 3.3 Effect of Individual Features on Creep Lifetime Determination in Specific Compositions

Based on the artificial Based on the AI model learned through the SHAP decision plot, the contribution of individual input features to the creep life determination process of specific test data can be explained. The evaluation is based on the standard composition (Table 4) with 730 h, which is the overall creep life of the learning data and calculates the effect of individual features on the SHAP value.

For example, in Fig. 6, the 100th data (0.12C-0.41Si-0.46Mn-12.1Cr-0.04Mo-0.01 W-0.23Ni-0.01 V-0.005Nb-0.013 N-0.0003B-0.05Co-0.817N\_Temp-1.145Test\_Temp)

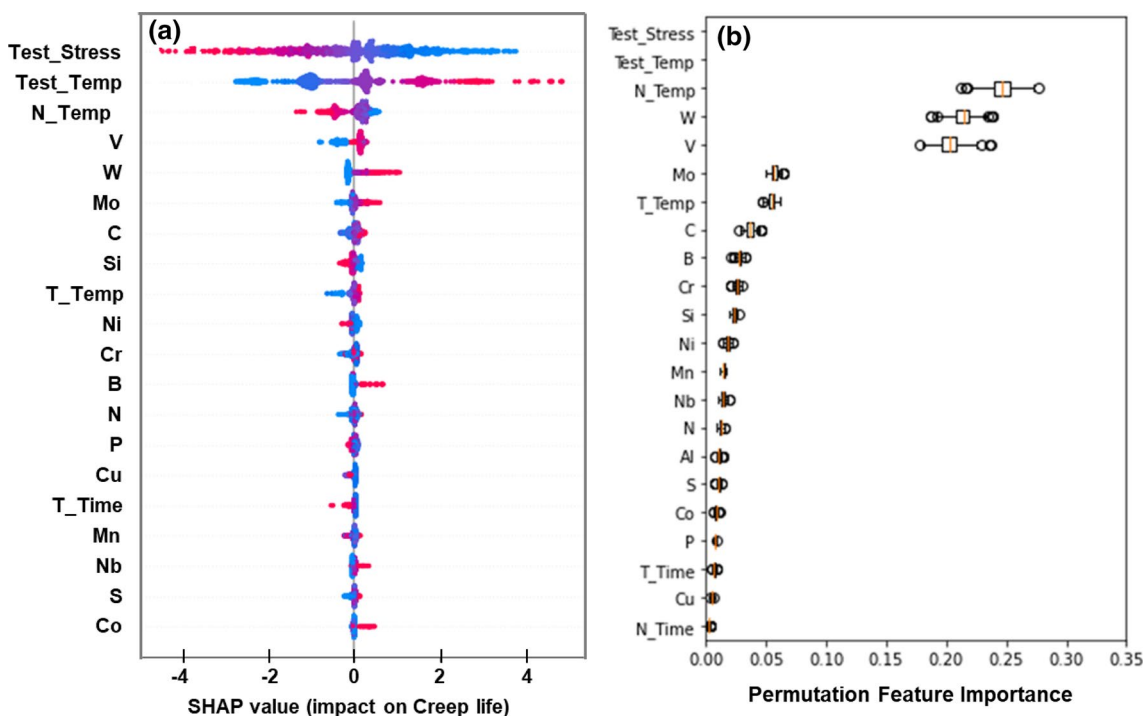
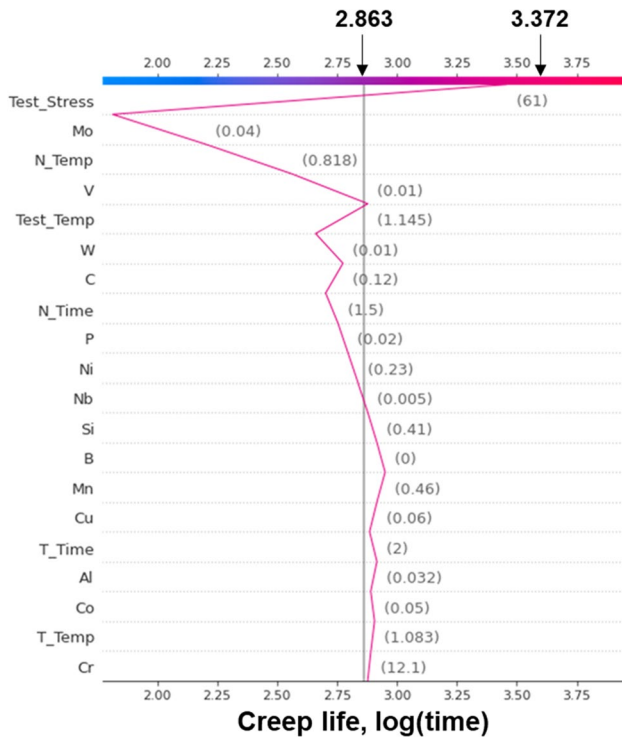


Fig. 5 A feature importance plot that can determine the influence of features on model prediction. a SHAP summary plot, b permutation importance plot

**Table 4** A standard alloy composition that provides a reference SHAP value

c	Si	Mn	P	s	Cr	Mo	W	Ni	Cu	V	Nb	N	Al	B	Co	N_Temp
0.05	0.19	0.56	0.011	0.016	11.86	0.27	0.46	0.48	0.22	0.26	0.06	0.042	0.045	0.012	1.336	0.77

**Fig. 6** SHAP decision plot showing the effect of individual features on the size of the 100th data value

consists of a combination of low Mo, V, and normalizing temperature, which decreases the SHAP value, and low Mn, high C, and low Test\_Stress, etc., which increases the SHAP value compared to the standard component. Through the analysis of these individual results, it is possible to explain why it has a creep life of 2,355 h (SHAP value: 3.372).

### 3.4 Analysis of the Influence of Alloying Elements on Creep Life Using XAI

As shown in the 22-dimensional data space (Fig. 2) including 16-dimensional components, there is a limit to applying the traditional regression technique to the prediction of the target value according to a specific feature change due to the multiple variance effect. Accordingly, the analysis results for six alloy elements with high feature importance indices are shown in Fig. 7 using the SHAP dependence scatter plot of the XAI technology based on the XGB model.

Figure 7 predicts the creep life according to the amount of change in alloying elements and is expressed as a relative SHAP value with the base SHAP value (730 h creep life) set to 0. The vertical spread of the SHAP values at a fixed composition is due to interaction effects with all the other variables in the figure. This means that the effect of multicollinearity among all features is reflected.

In the case of V (Fig. 7a), when less than 0.05 wt% is added, creep life becomes shorter than the standard creep life value of 730 h, and after increasing linearly to 0.1 wt%, creep life is saturated even if it increases to 0.25 wt%. W (b) and Mo (c) are predicted to continuously increase the creep life by a factor of 10 with the addition range. Using C (d), a creep life higher than the average can be obtained when at least 0.1% is added and receives a low feature importance rating due to the lower creep life increase effect compared to W and Mo. In the case of Si and Ni, however, it was predicted that the creep life would be reduced as the amount added increased. Therefore, the use of the currently proposed XAI technique can overcome the limitations of the existing experimental traditional method that can only consider a few types of alloy composition changes.

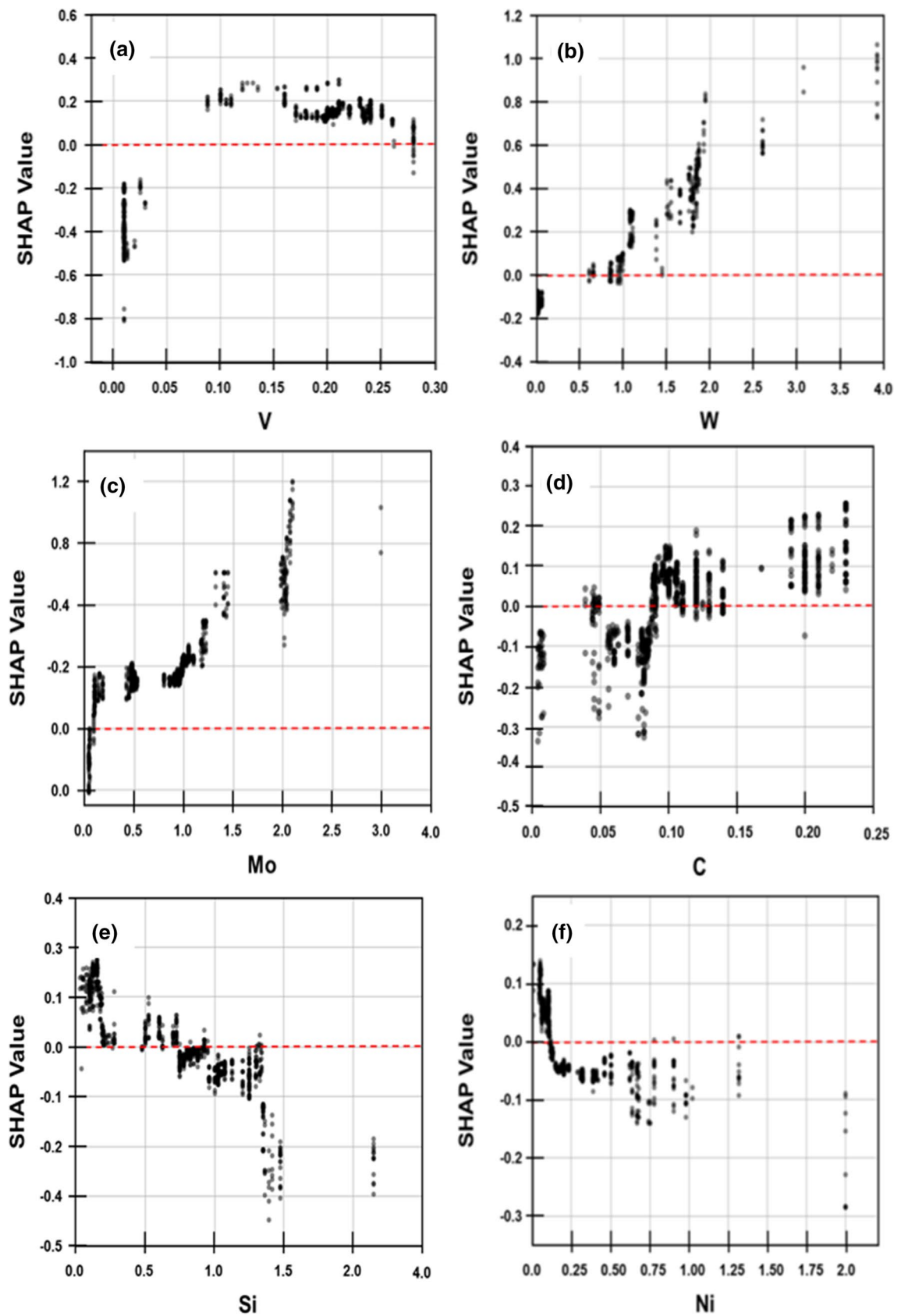
### 3.5 Optimization of Alloying Elements Using Genetic Algorithm

The effect of individual features on the target (creep life) can be analyzed through the SHAP dependence scatter plot in Sect. 3.4, but conversely, determining the optimal feature that satisfies a specific target with an experimental method is impossible due to the large number of variables that must be considered.

For example, even if 10 values are changed for each input feature,  $10^{21}$  combinations must be considered; therefore, the optimization process is very difficult. To solve this problem, in this work, the optimization process was performed using a GA. In this process, the values of the four main parameters were set as: crossover rate, 0.5; mutation rate, 0.1; population size, 200; and number of generations, 2000.

#### 3.5.1 Maximum Creep Life Prediction of the ASME SA213 T92 Material at an Operating Temperature of 650 °C and 100 MPa

The combination of the optimal feature with the longest creep life in the range that satisfies the alloy composition



**Fig. 7** SHAP dependence plot showing the change in SHAP Value (Creep life) value according to the addition of important alloying elements. **a** V, **b** W, **c** Mo, **d** C, **e** Si, and **f** Ni



**Table 5** Chemical composition range of ASME SA213 T92 code regulation

Element	C	Si	Mn	P	S	Cr	Mo	W	Ni	V	Nb	N	Al	B	N_Temp	T_Temp
Min	0.07	–	0.3	–	–	8.5	0.3	1.5	–	0.15	0.04	0.03	–	0.001	1040	730
Max	0.13	0.5	0.6	0.02	0.01	9.5	0.6	2.0	0.4	0.25	0.09	0.07	0.02	0.006	1080	800

**Table 6** The chemical composition with longest creep life in ASME SA213 T92 code regulation

	C	Si	Mn	P	S	Cr	Mo	W	Ni	V	Nb	N	Al	B	N_Temp	T_Temp	Max.Life
650 °C 100 MPa	0.09	0.015	0.326	0.007	0.01	9.10	0.59	1.99	0.01	0.22	0.085	0.06	0.001	0.006	1078	734	58,640

and heat-treatment specifications (Table 5) of ASME SA213 T92 was derived by applying the GA method based on the Ensemble 4 model (Table 6). The maximum creep life was predicted to be 58,640 h (4.76 SHAP value). From the point of view of domain knowledge, S has been treated as a harmful element because it reduces creep ductility [37]. In terms of creep life, S addition is predicted to increase creep life.

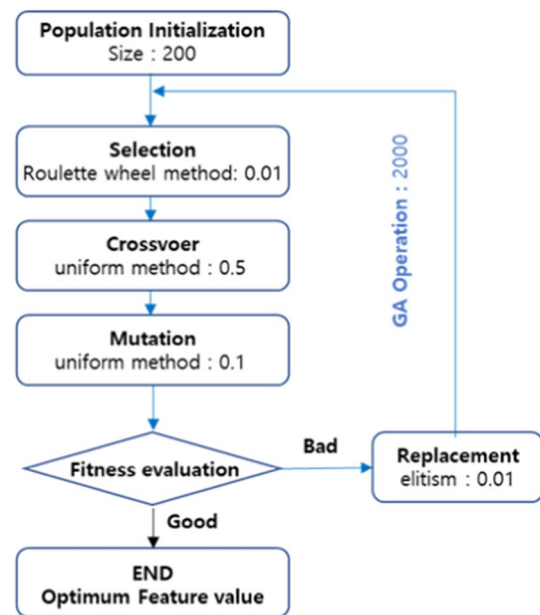
It shows that domain knowledge must be used in the process of securing raw data and subsequent data-preprocessing because the alloy must be designed in consideration of various requirements from the viewpoint of materials engineering. As a result of the GA model prediction, it is recommended that Si, Ni, and Al have as low a value as possible; therefore, it was considered necessary to lower the prescribed value in terms of creep life.

Figure 8 presents the result of change in creep life prediction according to changes in individual alloy composition based on the composition proposed to be optimal (Table 6).

In the case of W, it is predicted that the creep life increase effect starts to appear significantly from 1.8 wt% and for B from 0.003 wt%. Mo did not increase significantly, but the creep life continued to increase. Therefore, in the case of these alloy components, their maximum life is at the maximum allowable composition. By contrast, the maximum creep life was predicted at 0.23% for V, 0.09 wt% for C, and 9.10% for Cr. The lifespan can be rapidly reduced at 8.6% Cr or less.

### 3.5.2 Development of a New Alloy with Over 100,000 h Creep Life at an Operating Temperature of 650 °C and 100 MPa

Co, which is being considered by many researchers [38, 39], was selected for the development of a new alloy with over 100,000 h creep life at 650 °C and 100 MPa. As a result of predicting the optimum alloy composition having the maximum creep life using the GA model, it was confirmed that

**Fig. 8** Flow chart of genetic algorithm

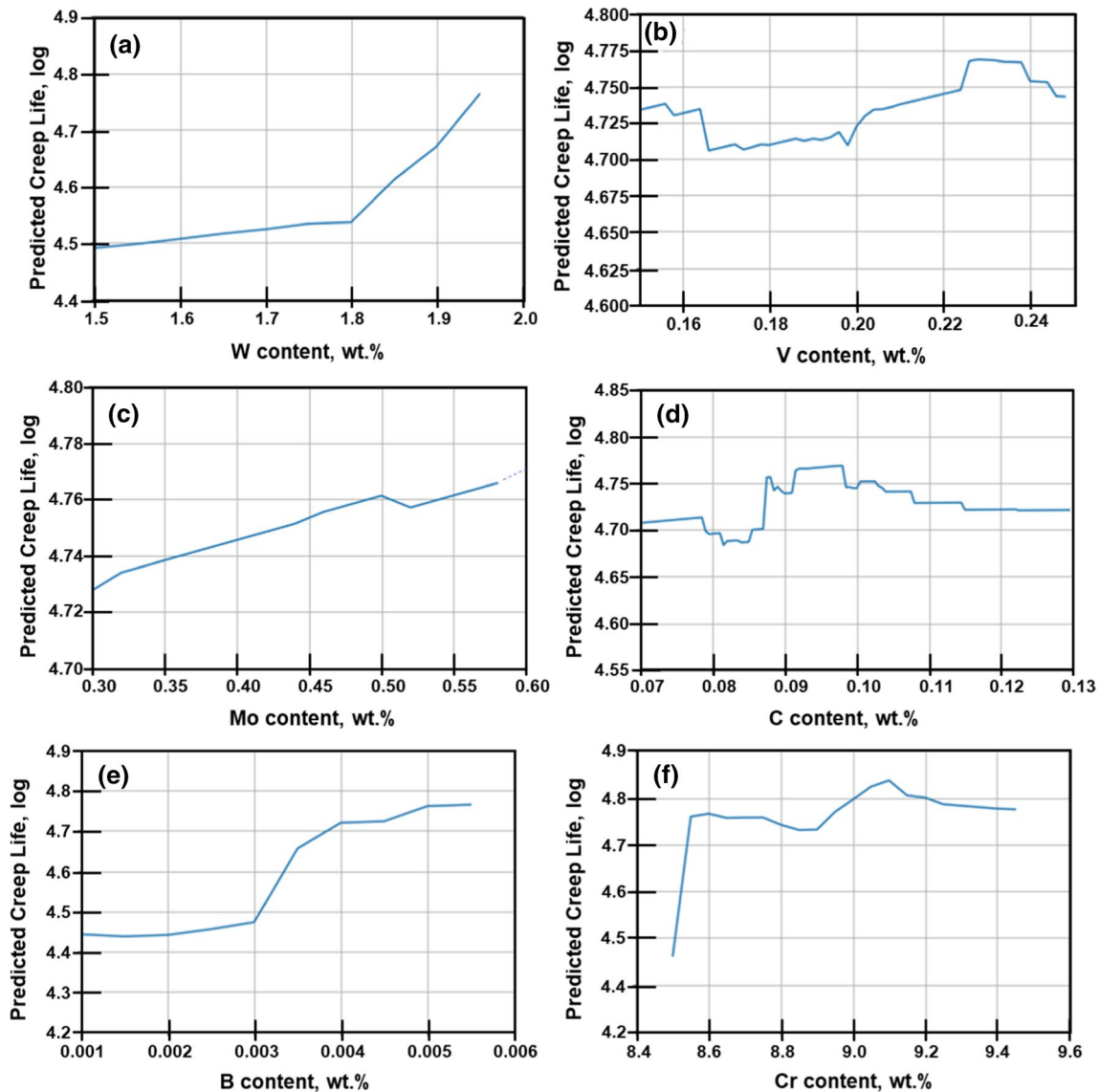
the chemical composition without Co addition in Table 6 as similar to the alloy composition in Table 7.

To evaluate the effect of Co addition on creep life, Cr was selected because it showed an interesting behavior in Fig. 8 and the analysis results using the GA model are shown in Fig. 9. Regardless of the addition of Co, the creep life trend according to the Cr content was predicted to be almost similar except near the maximum life, and it was found that the creep life was improved by about 40,000 h with addition of 2.5% Co. This tendency was also found for other alloying elements. That is, Co does not have a clear interaction with other alloying elements, and it could be determined that it strongly contributes to the solid solution-strengthening effect.

Figure 10 predicts the change in creep life according to the amount of Co added. As the amount of Co increases,

**Table 7** Proposed new alloy composition by applying the genetic algorithm

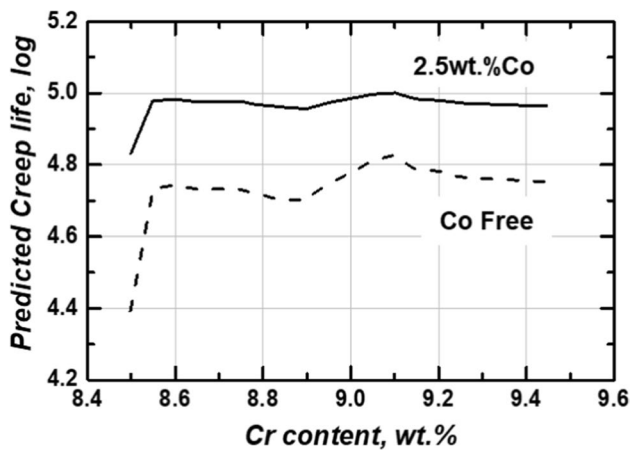
	C	Si	Mn	P	S	Cr	Mo	W	Ni	V	Nb	N	Al	B	Co	N_Temp	T_Temp	Max.Life
650 °C 100 MPa	0.09	0.03	0.31	0.015	0.01	9.10	0.59	1.99	0.01	0.23	0.07	0.06	0.001	0.006	2.5	1078	734	114,800

**Fig. 9** Creep life according to alloy element change predicted by Genetic algorithm based on Ensemble 4 model. **a** W, **b** V, **c** Mo, **d** C, **e** B and **f** Cr

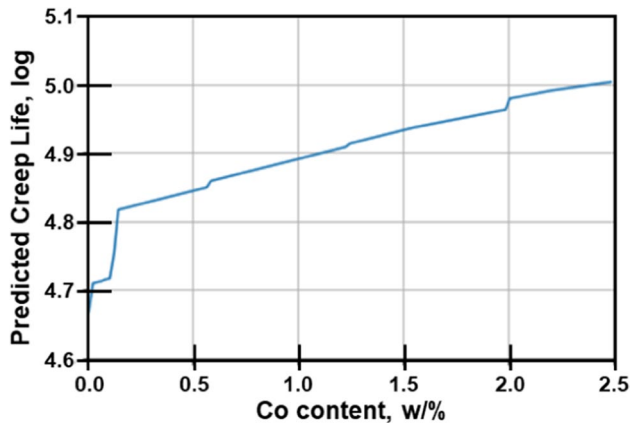
the creep life increases, and the maximum creep life was found to be 114,800 h. These results suggest the possibility of developing an alloy with a new composition that can be used with a creep life of 100,000 h under the operating conditions of 650 °C and 100 MPa (Fig. 11).

## 4 Conclusions

Among the various AI models, the XGB model was optimally selected as a single model, and the Ensemble 4 model consisting of four models (i.e., DNN, GB, XGB, and LGB) was optimally selected for the Ensemble 4 model. Based on this model, it was confirmed that creep



**Fig. 10** Analysis of the effect of Co on creep life according to the addition of Cr in the alloy of a new composition containing 2.5wt. Co



**Fig. 11** Prediction of creep life according to the addition of Co in the newly proposed alloy based on ASEM SA213 T92

life is affected in the order of normalization temperature, V, W, Mo, and C, as a result of feature importance evaluation using SHAP and premature values.

The existing life prediction model could not suggest an alloy composition with a desired creep life, but it could be confirmed that using the GA method an optimized alloy composition can be obtained.

For example, by adding Co to the specified composition range of SAME SA213 T92, a new alloy having a creep life of 100,00 h at 650 °C and 100 MPa having the following new composition was suggested: C, 0.09–Si, 0.03–Mn, 0.31–P, 0.015–S, 0.01–Cr, 9.10–Mo, 0.59–W, 1.99–Ni, 0.01–V, 0.23–Nb, 0.07–N, 0.06–Al, 0.001–B, 0.006–Co, 2.5. In the future, various AI technologies, such as XAI

technology, are expected to provide important information for the understanding of predictive models and the alloy development process, and for reducing the number of trial-and-error cycles for developers.

**Acknowledgements** The study was supported by the MOTIE Project No.20016472 and 20016054 of Research Center for Technological Innovation on Future Energy in Changwon National University.

## Declarations

**Conflict of interest** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

1. F. Masuyama, *ISIJ Int.* **41**, 612 (2001)
2. M. Subanović, J. Pirón, F. Zeller, M. Jarrar, A. Schneider, Development of a new high-performance martensitic heat-resistant steel for boiler applications, in *Proceedings of the ASME 2018 Symposium on Elevated Temperature Application of Materials for Fossil, Nuclear, and Petrochemical Industries*, Seattle, 3–5 April 2018, V001T01A008 (ASME, New York, 2018)
3. S. Yamasaki, M. Mitsuhashi, H. Nakashima, *ISIJ Int.* **58**, 1146 (2018)
4. K. Kimura, *Mater. High Temp.* **25**, 121 (2008)
5. K. Kimura, Creep rupture strength evaluation with region splitting by half yield, in *Proceedings of the ASME 2013 Pressure Vessels and Piping Conference*, Paris, 14–18 July 2013. vol. 6A: Materials and Fabrication, V06AT06A038 (ASME, New York, 2013)
6. M. Prager, *J. Press. Vessel Technol.* **117**, 95 (1995)
7. F.R. Larson, J. Miller, *Trans. ASME* **74**, 765 (1952)
8. F. Masuyama, *Int. J. Press. Ves. Pip.* **84**, 53 (2007)
9. A.K. Verma, J.A. Hawk, L.S. Bruckman, R.H. French, V. Romanov, J.L.W. Carter, *Metall. Mater. Trans. A* **50**, 3106 (2019)
10. H.K.D.H. Bhadesh, *ISIJ Int.* **39**, 966 (1999)
11. M. Liang, Z. Chan, Z. Wan, Y. Gan, E. Schlagen, B. Savija, *Cement Concrete Comp.* **125**, 104295 (2022)
12. J. Wang, Y. Fa, Y. Tian, X. Yu, *J. Mater. Res. Technol.* **13**, 635 (2021)
13. Y.-H. Sheu, *Front. Psychiatry* **11**, 551299 (2020)
14. D. Gunning, Explainable Artificial Intelligence (XAI), DARPA/I2O, Program Update November 2017
15. <https://ambiata.com/blog/2021-04-12-xai-part-1>
16. A. Mahendran, A. Vedaldi, Understanding deep image representations by inverting them, in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7–12 June 2015 (IEEE, New York, 2015), pp. 5188–5196
17. A. Dosovitskiy, T. Brox, Inverting visual representations with convolutional networks, in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27–30 June 2016 (IEEE, New York, 2016), pp. 4829–4837
18. B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, Learning deep features for discriminative localization, in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27–30 June 2016 (IEEE, New York, 2016), pp. 2921–2929
19. T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, 2nd edn. (Springer, New York, 2009), pp. 369–370

20. C. Molnar, *Interpretable Machine Learning* (Leanpub, Victoria, 2019), pp. 51–70
21. <https://www.oreilly.com/content/introduction-to-local-interpretable-model-agnostic-explanations-lime/>
22. <https://shap.readthedocs.io/en/latest/index.html>
23. M.L. Scott, S.-I. Lee, A unified approach to interpreting model predictions, in *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017)*, ed. by U. von Luxburg, I. Guyon, S. Bengio, H. Wallach, R. Fergus, Long Beach, 4–9 December 2017 (Curran Associates Inc., Red Hook, 2017)
24. L. Breiman, *Mach. Learn.* **45**, 5 (2001)
25. [https://scikit-learn.org/stable/modules/permutation\\_importance.html](https://scikit-learn.org/stable/modules/permutation_importance.html)
26. D. Opitz, R. Maclin, *J. Artif. Intell. Res.* **11**, 169 (1999)
27. R. Polikar, *IEEE Circ. Syst. Mag.* **6**, 21 (2006)
28. L. Rokach, *Artif. Intell. Rev.* **33**, 1 (2010)
29. H. Blockeel, Hypothesis space, in *Encyclopedia of Machine Learning*, ed. by C. Sammut, G.I. Webb (Springer, Boston, 2011), pp. 511–513
30. ASME BPVC. II. A-209 SA213/SA213M, Specification for seamless ferritic and austenitic alloy-steel boiler, superheater, and heat exchanger tubes, pp. 285–299 (2019)
31. <https://www.phase-trans.msm.cam.ac.uk/map/map.html>
32. B.O. Kong, S.H. Ryu, Y.S. Lee, J.I. Suk, J.T. Kim, Y.S. Yoo, Application of neural network data modeling for estimating of material properties and creep life, in *Proceedings of the 17th Conference on Mechanical Behaviors of Materials*, ed. by J.H. Hong, Y.S. Kim. Pohang, 14 November 2003 (The Korean Institute of Metals and Materials, Seoul, 2003), pp. 300–307
33. M. Mitchell, *An Introduction to Genetic Algorithms* (MIT Press, Cambridge, 1996)
34. [https://en.wikipedia.org/wiki/Genetic\\_algorithm](https://en.wikipedia.org/wiki/Genetic_algorithm)
35. H. Zhou, Z. Deng, Y. Xia, M. Fu, *Neurocomputing* **216**, 208 (2016)
36. J.T. Terpstra, C.-H. Chang, R.C. Magel, *J. Stat. Comput. Simulat.* **81**, 1381 (2011)
37. J. Parker, J. Siefert, *Adv. Mater. Sci. Eng.* **2018**, 6789563 (2018)
38. A. Kipelova, M. Odnobokova, A. Belyakov, R. Kaibyshev, *Metall. Mater. Trans. A* **44**, 577 (2013)
39. L. Helis, Y. Toda, T. Hara, H. Miyazaki, F. Abe, *Mater. Sci. Eng. A* **510–511**, 88 (2009)

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.