Predicting solidification cracking susceptibility of stainless steels using machine learning

S Feng and H B Dong
Department of Engineering, University of Leicester, UK
E-mail: hd38@le.ac.uk

Abstract. Machine learning, which reveals the complex nonlinear relationship in the archived data, is a powerful complement to theory, experiment, and modeling. In this study we attempted to predict solidification cracking susceptibility of stainless steels as a function of chemistry and processing parameters using machine learning with a data set that contains about 600 longitudinal varestraint test results. Four machine learning models, i.e. decision tree, random forest, shallow neural network and deep neural network, were used to mine the data set. Our results show: deep neural network outperformed other models in prediction accuracy; tree-based models have accepted accuracy and better interpretability than neural network; machine learning models transforms scattered experimental data points into a map in high-dimensional chemistry and processing parameters space. The combination of different machine learning models reveals that the solidification cracking susceptibility of stainless steels was mainly determined by the ratio of Ni content to Cr content, impurity element content and the strain level.

1. Introduction
Solidification cracking is one of the most serious defects which widely occurs in processes involve liquid to solid transformation like casting, welding, and additive manufacturing. It occurs at the last stage of solidification when liquid films exist between dendrites boundaries where local strains cannot be accommodated by liquid feeding and solid deformation [1]. It's a complex problem because of interaction of material factors, mechanical restraint factors, thermal factors in the final crack nucleation and propagation [2]. Metallic material processability e.g. weldability and castability is closely linked with its solidification cracking susceptibility. Due to its complexity and importance, solidification cracking has been researched widely with experiments, theories, and modelling. Studying solidification cracking is time-consuming and expensive. Theories about solidification cracking susceptibility based either on metallurgical features, or strain, or strain rate, or stress criteria are either over-simplified or lack physical basis, thus preventing them from practical application as no existing theory can include all three types of factors mentioned above [3]. Successful computer numeric models which take into consideration of thermal and mechanical interaction, intergranular network geometry and flow field can only predict the position of most susceptible to solidification cracking but not alloys solidification cracking susceptibility [4]. To further study this type multiple variables nonlinear problem, data-based and data-driven machine learning methods, e.g. neural network and tree-based models, increasingly become an alternative to physically based analytical and numerical methods [5]. Based on reliable experimental data, machine learning methods can depict alloy solidification cracking susceptibility as function of alloy composition and processing parameters in a high-dimensional space, like that was shown in our previous work [6].
To explore the relationship between solidification cracking susceptibility, composition and processing parameters, intermediate steps (bridges), i.e. atomic/meso/micro/macro structure analysis, are commonly required in traditional material science route. The relationship between composition, processing parameters and structure and the relationship between structure and solidification cracking susceptibility are researched separately first. A lot of new terminologies and models are proposed like ratio of chromium equivalent to nickel equivalent, primary ferritic/austenitic solidification mode, impurities content, solidification region, etc. to simplify analysis that results in an indirect relationship between solidification cracking susceptibility and three main influential factors. But the main disadvantages of traditional material science methods are the relationship found is always limited to specific composition and processing parameters and most of them are qualitative descriptions, a universal relationship is rarely found, especially for the complex nonlinear phenomena. This shortcoming could be overcome by many data-based methods especially the deep learning by which the hierarchy structure of data can be features engineered automatically [7,8], thus we can get solidification cracking susceptibility as function of composition and processing parameters, i.e. quantitative mappings from inputs to outputs, directly without the aids of structure analysis.

In this paper we combined different machine learning models including deep neural network of high accuracy and tree-based models of high interpretability to predict solidification cracking susceptibility of stainless steels.

2. Methods

2.1. Data collection

Dozens of test methods exist in order to qualitative and quantitatively evaluate the solidification cracking susceptibility. A controlled strain applied on a geometrically simple specimen and test method of good repetition is preferred for evaluation of cracking tendency such as longitudinal varestraint test, transverse varestraint test, etc. In longitudinal varestraint test a definite bending strain is applied on weld of sheet metals specimen, and total crack length (TCL) and maximum crack length, brittle temperate range are measured as the indicators for solidification cracking susceptibility. A large amount of longitudinal varestraint test data are available in literature. That is why we chose longitudinal varestraint test data in published literature to compile our solidification cracking susceptibility data set. In this study, a data set containing 600 longitudinal varestraint test results of stainless steels was collected from literature. The data set contains information of ferritic, austenitic, duplex, precipitation hardening wrought and casting stainless steels, and comprises 4 types information: stainless steels composition information of 16 elements (contents of C, Si, Mn, P, S, Cr, Ni, Mo, N, Nb, Co, Cu, Al, Ti, V, B, all compositions are expressed in wt% in this paper), samples thickness (Th, mm) and 3 welding parameters (welding current: I, A; voltage: U, V; velocity: Ve, mm/s), restraint information (applied strain: ε, %) and solidification cracking susceptibility indicator total crack length (TCL, mm). The elements contents which were not specified in literature were assumed as follow: Al = 0.02, N = 0.02, Mo = 0, Nb = 0, Cu = 0, V = 0, B = 0. Data missing too much information was discarded, and the final size of training/testing data set is 487.

2.2. Neural network’s preprocess and training algorithm

The input variables X were normalized between -0.5 and +0.5, the target variables Z were normalized between -0.99 and +0.99. The neural networks used in this study consisted of 21 input neurons (one neuron corresponds to one input variables), one to several hidden layers, a few hidden neurons (the neurons number is variable) in each hidden layer, and one output neurons for TCL. The transfer function of hidden layer is hyperbolic tangent function \( \tanh(x) \). The transfer function of output layer is linear function \( y = x \).

To achieve unbiased training/testing data division, all data were divided into groups first, and in a group only one or two variables vary, then training data and testing data were randomly chosen in a 2 to 1 ratio, i.e. 2/3 data set (324 samples) was used for training, and the remaining 1/3 data set (163
samples) were kept unseen by neural network in training to test the generalization performance (prediction accuracy on unseen data set) of the trained neural network.

To improve generalization and reduce the possibility of overfitting, a type of regularized performance function \( \text{mse}_{\text{reg}} \), which combines mean square error with mean square weight, was used in training shallow neural network and fine-tuning deep neural network (after pre-training). The weights and biases (special type of weights with input value of 1) are determined through training neural network which involved minimization of the regularized performance function through gradient descent algorithm (backpropagation algorithm) and auto determination of optimal performance ratio \( \gamma \) using Bayesian regularization proposed by David MacKay. The training and regularization details are described elsewhere [9, 10].

One hidden layer shallow neural network is a mature tool for science and engineering field, including material science. The commonly used ways to increase prediction accuracy of a shallow neural network include: increasing data set size; increasing variables or features of data set; increasing neuron number in hidden layer; changing network’s initial weights and biases values; increasing training epoch. And the ways to increase prediction accuracy of the deep neural network include: increasing hidden layer number; changing neuron numbers in hidden layers; pretraining deep neural network using restrict Boltzmann machine and auto-encoder to achieve optimal initiation; changing network structure (from full connected to partially connected); changing training algorithms and using special training strategy (e.g. drop out) [11, 12]. In this study, we tried shallow neural networks of neuron number from 1 to 35 and train every shallow neural network configuration more than 100 times using different initiation values, then the optimal trained shallow neural network was chosen to make prediction. We also attempted 3 hidden layers deep neural networks of 21-(4-3-2)-1 and 21-(5-4-3)-1 structure, 4 hidden layers deep neural networks of 21-(5-4-3-3)-1 and 21-(6-5-4-3)-1 structure, and 5 hidden layers deep neural networks of 21-(6-5-4-3-3)-1 and 21-(5-4-3-3-3)-1 structure. The numbers in bracket represent neuron numbers of hidden layers in sequence. A selection of initiation values for stacked auto-encoder were used to generate different initiation conditions for deep neural networks. Every deep neural network configuration was trained more than 100 times using different initiation conditions, then the optimal deep neural networks was chosen to make prediction. After training Pearson correlation coefficient \( R \) of target values and neural network prediction values was calculated as the index of training / testing accuracy.

### 2.3. Tree-based machine learning models

Tree-based machine learning models are a family of non-parametric, supervised methods. The foundation of tree-based models is the decision tree wherein a series of decision rules (e.g., “If \( A > B \)” ) are linked. Random forest model is the ensemble of decision trees. A decision tree looks like an upside-down tree, with the first decision rule at the top and subsequent decision rules spreading out below. In a decision tree, every decision rule occurs at a decision node, with the rule creating branches leading to new nodes. A branch without a decision rule at the end is called a leaf. Decision tree regression models attempt to find a decision rule that produces the greatest decrease mean squared error at a node. One benefit of tree-based models is their interpretability. Visualization of decision trees can give researchers an intuitive knowledge of feature importance.

In this work, all machine models are realized on MATLAB 2018a with its statistics and machine learning toolbox and neural network toolbox.

### 3. Results

#### 3.1. Training and testing accuracy

Table 1 shows the accuracies of four machine learning models on testing data set. Random forest’s accuracy (0.88) is better than that of a single decision tree (0.75) and is close to that of shallow neural network. Deep neural network has the highest prediction accuracy in the four machine learning models.
Table 1. The best testing accuracies of four machine learning models

|                      | Decision tree | Random forest | Shallow neural network | Deep neural network |
|----------------------|---------------|---------------|------------------------|---------------------|
| Testing accuracy     | 0.75          | 0.88          | 0.89                   | 0.93                |

3.2. Neural network structure

The structure schematics of the optimal shallow and deep neural network are shown in figure 1. It shows the optimal 21-(6-5-4-3)-1 deep neural network has more compact structure (the hidden layer of deep neural network is narrower than that of shallow neural network) and less parameters (210 vs. 484) than the optimal 21-(21)-1 shallow neural network. The trained neural network can be easily used to estimate solidification cracking susceptibility dependence on compositions and strains (small strains can be utilized to predict solidification cracking susceptibility in casting and large strains can be chosen to predict solidification cracking susceptibility in welding and additive manufacturing) when the time consuming and expensive experimental results are not available. The following sections provide several examples of application, but the application of the machine learning models is not limited to those.

![Figure 1](image1.png)

(a) shallow neural network

![Figure 1](image2.png)

(b) deep neural network

Figure 1. (a) The optimal shallow neural network’s structure schematic; (b) the optimal deep neural network’s structure schematic.

4. Predicting solidification cracking susceptibility

4.1. Feature importance

Decision tree can give us an intuitive knowledge of how solidification cracking susceptibility is calculated according to a serial of rules. Figure 2 visualizes a decision tree used for predicting solidification cracking susceptibility. The Ni content is used as the first decision rule and the strain is used as the second decision rule, this shows the importance of the Ni content and strain in determining stainless steels’ solidification cracking susceptibility.

![Figure 2](image3.png)

Figure 2. A decision tree for predicting solidification cracking susceptibility.
Besides the intuitive impression of feature importance, tree-based machine learning models can give the value of feature importance. Figure 3 is the feature importance for solidification cracking susceptibility derived from the optimal random forest model. Major alloy elements Ni and Cr, minor alloy elements Mn and Si, impurity elements C, N, P, S, and the strain applied on the specimen play the most important role in determining the solidification cracking susceptibility of stainless steels. The knowledge derived can be used to guide next step predictions. A serial of predictions about solidification cracking susceptibility dependence on compositions and processing parameters in the following part will give us more details and deeper understanding of this complex problem.

![Figure 3](image)

**Figure 3.** Feature importance for solidification cracking susceptibility from the random forest model.

![Figure 4](image)

**Figure 4.** The prediction of 8 stainless steels’ solidification cracking susceptibility dependence on Ni and Cr (their Ni and Cr contents are in AISI specification). Compositions are shown in table 2.

Processing parameters: Th=3.18 mm, I=100 A, U=12 V, Ve=4.23 mm/s, strain=3%.

4.2. Solidification cracking susceptibility dependence on Ni and Cr

The prediction of 8 stainless steels’ solidification cracking susceptibility dependence on the main alloy elements Ni and Cr is shown in figure 4. The predictions are in good accordance with metallurgical experience: fully austenitic stainless steels like AISI 316 is more susceptible to solidification cracking than stainless steels contain a certain amount of ferrite like AISI 304; C which tend to segregate to grains boundaries increase cracking susceptibility (see the contour difference between AISI 316L and AISI 316); machine learning regression has transformed scattered data points from our data set into an expressive high-dimensional map and those contours are some slices of it. The tendency of
increasing solidification cracking susceptibility with the increment of Ni is obviously illustrated in the figure 4. High ratio of Cr to Ni (or Cr equivalent to Ni equivalent) is good for solidification cracking resistance by forming a certain amount of ferrite like that in AISI 301 and 304 etc., because ferrite can accept more impurity elements (like S and P) than austenite and irregular ferrite/austenite grain boundary is not in favour of the propagation of cracking cracks.

**Table 2.** Compositions (wt%) of stainless steels used in predictions in figure 4 (varied Ni and Cr).

| Code | C  | Si | Mn | P   | S   | Cr    | Ni | Mo | N  | Nb | Cu | Al | Ti |
|------|----|----|----|-----|-----|-------|----|----|----|----|----|----|----|
| 301  | 0.10 | 1.0 | 2.0 | 0.02 | 0.02 | 16.0-18.0 | 6.0-8.0 | -  | 0.10 | -  | -  | -  | -  |
| 301L | 0.03 | 1.0 | 2.0 | 0.02 | 0.02 | 16.0-18.0 | 6.0-8.0 | -  | 0.10 | -  | -  | -  | -  |
| 304  | 0.08 | 1.0 | 2.0 | 0.02 | 0.02 | 18.0-20.0 | 8.0-11.0 | -  | 0.10 | -  | -  | -  | -  |
| 304L | 0.03 | 1.0 | 2.0 | 0.02 | 0.02 | 18.0-20.0 | 8.0-12.0 | -  | 0.10 | -  | -  | -  | -  |
| 316  | 0.08 | 0.75 | 2.0 | 0.02 | 0.02 | 16.0-18.0 | 10.0-14.0 | 2.0 | 0.10 | -  | -  | -  | -  |
| 316L | 0.03 | 0.75 | 2.0 | 0.02 | 0.02 | 16.0-18.0 | 10.0-14.0 | 2.0 | 0.10 | -  | -  | -  | -  |
| 321  | 0.08 | 0.75 | 2.0 | 0.02 | 0.02 | 17.0-19.0 | 9.0-12.0 | -  | -  | -  | -  | -  | 0.5 |
| 347  | 0.08 | 0.75 | 2.0 | 0.02 | 0.02 | 17.0-19.0 | 9.0-13.0 | -  | -  | 0.5 | -  | -  | -  |

4.3. Solidification cracking susceptibility dependence on P and S

The prediction of 8 stainless steels’ solidification cracking susceptibility dependence on impurity elements P and S is shown in figure 5. The tendency of increasing solidification cracking susceptibility with the increment of P and S is shown in the prediction. The extent of increasing susceptibility with P and S are different among different steels. The solidification cracking susceptibility of AISI 316 is very sensitive to the P and S contents: small difference in P and S contents can show great difference in solidification cracking behaviour. But that of AISI 316L with less C content is less sensitive to the P and S contents. Figure 5 also shows that impurity elements P and S have different impact on solidification cracking susceptibility (most contours are not symmetry along P=S diagonal), the common assumption of linear superposition P and S used in estimating solidification cracking susceptibility is not always proper.

**Figure 5.** The prediction of 8 stainless steels’ solidification cracking susceptibility dependence on P and S. Compositions are shown in table 3. Processing parameters: Th=3.18 mm, I=100 A, U=12 V, Ve=4.23 mm/s, strain=3%.
The tendency of increasing solidification cracking susceptibility with the increase of AISI 316 is very sensitive to the C and N contents: small difference in C and N contents can show great difference in solidification cracking behaviour, but cracking susceptibility of AISI 301 and 304 etc. are less sensitive.

4.4. Solidification cracking susceptibility dependence on C and N

The prediction of 4 stainless steels’ solidification cracking susceptibility dependence on C and N is shown in figure 6. The tendency of increasing solidification cracking susceptibility with the increasement of C and N is revealed in the prediction. The extent of increasing susceptibility with C and N varied among different steels due to the different alloy elements in corresponding steels. The solidification cracking susceptibility of AISI 316 is very sensitive to the C and N contents: small difference in C and N contents can show great difference in solidification cracking behaviour, but cracking susceptibility of AISI 301 and 304 etc. are less sensitive.

| Code | Si  | Mn  | P   | S   | Cr  | Ni  | Mo  | N   | Nb  | Cu  | Al  | Ti |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 301  | 0.10| 1.0 | 2.0 | 0.001-0.040 | 0.001-0.040 | 17.0 | 7.0 | -   | 0.10 | -   | -   | -   |
| 301L | 0.03| 1.0 | 2.0 | 0.001-0.040 | 0.001-0.040 | 17.0 | 7.0 | -   | 0.10 | -   | -   | -   |
| 304  | 0.08| 1.0 | 2.0 | 0.001-0.040 | 0.001-0.040 | 19.0 | 9.5 | -   | 0.10 | -   | -   | -   |
| 304L | 0.03| 1.0 | 2.0 | 0.001-0.040 | 0.001-0.040 | 19.0 | 9.5 | -   | 0.10 | -   | -   | -   |
| 316  | 0.08| 0.75| 2.0 | 0.001-0.040 | 0.001-0.040 | 17.0 | 12.0| 2.0 | 0.10 | -   | -   | -   |
| 316L | 0.03| 0.75| 2.0 | 0.001-0.040 | 0.001-0.040 | 17.0 | 12.0| 2.0 | 0.10 | -   | -   | -   |
| 321  | 0.08| 0.75| 2.0 | 0.001-0.040 | 0.001-0.040 | 18.0 | 9.5 | -   | -   | -   | -   | 0.5 |
| 347  | 0.08| 0.75| 2.0 | 0.001-0.040 | 0.001-0.040 | 18.0 | 9.5 | -   | 0.5  | -   | -   | -   |

4.4. Solidification cracking susceptibility dependence on strain

The prediction of AISI 316 stainless steel’s solidification cracking susceptibility dependence on strain is shown in figure 7. The tendency of increasing solidification cracking susceptibility with the increment of strain is correctly revealed in the prediction. This tendency is also a nonlinear behaviour, which is very difficult for physical models to depict, but is quite easy for machine learning models.

Figure 6. The prediction of 4 stainless steels’ solidification cracking susceptibility dependence on C and N. Compositions are shown in table 4. Processing parameters: Th=3.18 mm, I=100 A, U=12 V, Ve=4.23 mm/s, strain=3%.

Table 4. Compositions (wt%) of the 4 stainless steels used in predictions in figure 6 (varied C and N).

| Code | C   | Si  | Mn  | P   | S   | Cr  | Ni  | Mo  | N   | Nb  | Cu  | Al  | Ti |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 301  | 0.001-0.15| 1.0 | 2.0 | 0.02 | 0.02 | 17.0 | 7.0 | -   | 0.001-0.15| -   | -   | -   | -   |
| 304  | 0.001-0.15| 1.0 | 2.0 | 0.02 | 0.02 | 19.0 | 9.5 | -   | 0.001-0.15| -   | -   | -   | -   |
| 316  | 0.001-0.15| 0.75| 2.0 | 0.02 | 0.02 | 17.0 | 12.0| 2.0 | 0.001-0.15| -   | -   | -   | -   |
| 321  | 0.001-0.15| 0.75| 2.0 | 0.02 | 0.02 | 18.0 | 9.5 | -   | 0.001-0.15| -   | -   | -   | 0.5 |

4.5. Solidification cracking susceptibility dependence on strain

The prediction of AISI 316 stainless steel’s solidification cracking susceptibility dependence on strain is shown in figure 7. The tendency of increasing solidification cracking susceptibility with the increment of strain is correctly revealed in the prediction. This tendency is also a nonlinear behaviour, which is very difficult for physical models to depict, but is quite easy for machine learning models.
Figure 7. The prediction of AISI 316 stainless steel’s solidification cracking susceptibility dependence on strain with variations in C and N. Compositions are shown in table 4. Processing parameters: Th=3.18 mm, I=100 A, U=12 V, Ve=4.23 mm/s.

5. Conclusions
Four machine learning models, i.e. shallow neural network, deep neural network, decision tree and random forest, were used to fit solidification cracking susceptibility of stainless steels as a function of chemistry and processing parameters. Neural network models outperformed tree-based models in prediction accuracy; tree-based models have accepted accuracy and better interpretability than neural networks; the optimal deep neural network has more compact structure and less parameters than the optimal shallow neural network.

Through machine learning regression, vast scattered non-systematic experimental data buried in literature can produce simple quantitative expression: specific material property as function of chemistry composition and processing parameters, etc. The derived models can be used in material property prediction, new alloys development, comparison with experimental results, etc.

The combination of different machine learning models reveals that the solidification cracking susceptibility of stainless steels was mainly determined by Ni content / Cr content (or ferrite content / austenite content), impurity content and the strain level.

Acknowledgements
Shuo Feng wishes to acknowledge EPSRC CDT (Grant No: EP/L016206/1) in Innovative Metal Processing for providing a PhD studentship for this study.

Reference
[1] Aucott L, Huang D, Dong H B, Wen S W, Marsden J, Rack A and Cocks A C F 2018 Metall. Mater. Trans. A 49 1674–82
[2] Eskin D G, Suyitno and Katgerman L 2004 Prog. Mater. Sci. 49 629–711
[3] Kou S 2015 Acta Mater. 88 366–74
[4] Bordreuil C and Niel A 2014 Comput. Mater. Sci. 82 442–50
[5] Mater A C and Coote M L 2019 J. Chem. Inf. Model. 59 2545–59
[6] Feng S, Zhou H and Dong H 2019 Mater. Des. 162
[7] Butler K T, Davies D W, Cartwright H, Isayev O and Walsh A 2018 Nature 559 547–55
[8] LeCun Y, Bengio Y and Hinton G 2015 Nature 521 436–44
[9] MacKay D J C 1992 Neural Comput. 4 448–72
[10] Bhadeshia H K D H, Dimitriu R C, Forsik S, Pak J H and Ryu J H 2009 Mater. Sci. Technol. 25 504–10
[11] Demuth H B, Beale M H, De Jess O and Hagan M T 2014 Neural network design (Martin Hagan)
[12] Goodfellow I, Bengio Y and Courville A 2016 Deep learning (MIT press)