Article

Machine learning predictions of additively manufactured alloy crack susceptibilities

Shamak Gowda¹*, Chloe Hung²*, Alice Zou³*, Kaitlyn Mullin⁴, Caelin Muir⁴

¹ Basis Independent Fremont, Fremont, California
² Monta Vista High School, Cupertino, California
³ Valley Christian High School, San Jose, California
⁴ Department of Materials, University of California, Santa Barbara, Santa Barbara, California
*These authors contributed equally to this work

SUMMARY

Additive manufacturing (AM) is revolutionizing the production of complex metal parts, offering both extreme precision and cost reduction. However, structural defects such as internal cracking can pose serious issues in AM parts, especially in safety-critical applications such as aerospace and automotive production. Empirically testing or running high-throughput simulations on a high volume of alloy formulas to quantify these cracking susceptibilities is overly costly and time-consuming. Recent developments in machine learning (ML) offer a more efficient solution to predicting the properties of a high volume of materials without the time cost needed to run mechanical simulations. In this study, we analyzed the solidification cracking susceptibilities (SCS) of potential printing alloys in safety-critical, additively manufactured parts. Based on the link between SCS and secondary alloy properties, we hypothesized that a multi-model ML pipeline would predict an alloy's solidification cracking susceptibility at a higher accuracy than a model trained on alloy composition alone. Since predicting SCS directly from alloy composition eliminates the need for empirical measurements, our pipeline also predicted these secondary attributes directly from the original composition. We found that a pipeline using Random Forest architecture models performed at the highest accuracy and precision, outperforming the precision of a random forest model with no secondary predictions. Furthermore, the results of this study indicated a correlation between SCS and the included secondary attributes, which can be investigated in further research, as these attributes have not yet been factored into current SCS quantification methods.

INTRODUCTION

Additive manufacturing (AM) uses novel techniques to construct 3D objects directly from computer-aided design (CAD) models (1). Differing from traditional manufacturing techniques such as casting and forging, AM processes offer complex geometries and designs, while also reducing part count (2,3). AM has allowed for the manufacture of metals that have traditionally been extremely challenging

due to immense strength, high melting points, and inability to maintain integrity in contact with carbon, oxygen, and nitrogen (4). In 2023, Relativity Space achieved a significant milestone with the launch of Terran 1, a small satellite launch vehicle that showcased the cutting-edge capabilities of metal AM (5). Through 3D printing, Terran 1 dramatically reduced manufacturing time from traditional manufacturing, enabling faster turnaround and increased production efficiency while maintaining the required precision and quality standards (5,6). One notable feature of its design was its nine Aeon engines, printed from a proprietary copper alloy engineered to withstand the extreme heat generated during propellant combustion (5). Examples such as Terran 1 demonstrate how AM has become prevalent in greater applications due to its rapid assembly and broader access to previously exclusive technologies such as rocket engines and parts that require intricate meshes and geometries (2). Furthermore, industries such as medicine, aeronautics, and aerospace have recognized the potential of AM to create lightweight, strong, and customized parts (2).

In safety-critical applications, it is crucial that alloys maintain their durability even under extreme conditions, such as high temperatures and velocities. Several such possible applications of AM include nuclear fusion, hypersonic leading edges, gas/aircraft turbine engines, and atmospheric re-entry (4). However, one barrier to using AM over traditional manufacturing is the need to meet certain industry standards on material performance (3). For example, aerospace manufacturers must prove that their AM engine parts have a high fatigue resistance for flight suitability to mitigate the risk of part failure (3). In the case of laser powder bed fusion (7), the alloy composition directly affects printability and the presence of defects in the finished part, making research and discovery of optimal compositions critically important. An example of this can be found in the propensity for cracking during solidification, known as the solidification cracking susceptibility (SCS) of a printed material (8). These defects occur when elements begin to unevenly distribute inside the melt pool as the material is solidifying. The change in liquid composition from the start to the end of solidification causes a difference in local melting point. At the same time, the portion of the melt pool that has solidified has contracted. If there is a large difference in temperature between the start and end of solidification, shrinkage can occur and stop proper liquid feeding, causing the formation of a crack. (6,9). Following solidification, additional solid-state cracking may occur if the materials do not have enough ductility to overcome the

stress from rapid shrinking and cooling (10). The propensity for cracking during solidification is known as the solidification cracking susceptibility (SCS) of a printed material. A higher value of SCS indicates higher susceptibility to cracking upon solidification which would present potential weaknesses in manufactured parts. Currently, the computation of SCS involves temperature and mechanical property measurements at specific solid fraction values, which require ultrahigh temperature capabilities (11). T.W. Clyne's definition of SCS can be calculated using Equation 1, with fs representing the solid fraction, T denoting temperature, t_v denoting vulnerable period, and t_r denoting time available for stress release. (11). Traditionally, the temperatures required for this calculation are obtained through a series of melting experiments, or thermodynamic simulation approaches, such as CALPHAD, which is the approach utilized by ThermoCalc (20). While these conventional calculation methods yield accurate results, they are expensive and time-consuming, as experimental tests and simulations can take days or even years to accurately estimate an alloy's SCS and other properties (12). These limitations generate the need for techniques capable of leveraging data-driven analysis to achieve faster and more

$$Clyne\ SCS\ =\ \frac{t_v}{t_r} = \frac{\frac{T_{fs=0.90}-T_{fs=0.99}}{R}}{\frac{T_{fs=0.40}-T_{fs=0.90}}{R}} \quad \text{(Equation 1)}$$

cost and time-efficient predictions, such as ML (13).

Within the last decade, there has been an increasing interest in employing ML to predict the mechanical properties of top-performing additive manufacturing alloys (14). The critical nature of safety for these components further motivates the need to develop reliable ML predictions for assessing structural integrity. Earlier research has focused on aluminum-based alloys, specifically eutectic Al-Si-based alloys, but due to inconsistencies in microstructural characteristics, such as grain size and geometry of boundary layers, of aluminum under directional solidification, more progress has been made with Ti- and Ni-based alloys (15). However, not all of these alloys are printable because the interaction of solid and liquid during solidification in printing environments can lead to extensive cracking. Limited research has been conducted in extending ML approaches to consider the defect resistance of an alloy during the printing process (15). Previous research has proposed the use of shallow neural networks (SNNs), deep neural networks (DNNs), and support vector regression (SVR) to calculate changes in SCS due to changes in two compositions of constituent stainless steels (12). However, this study used a smaller dataset of under 500 data points and exclusively focused on a specific subset of alloys. Moreover, their approach required longitudinal strain test data as input for predictions, again requiring manufacturers to obtain and experimentally manipulate the metal (12). We extended their methodology to encompass a broader range of alloys, employed a much larger dataset, and eliminated the need for any physical testing data as input.

We hypothesized that adding feature inputs of ductility, solidus, and general stacking fault energy (GSF) would result in a more accurate estimate of Clyne SCS than a model without these features. In this study, we present a novel method capable of predicting SCS solely from a given alloy composition. We accomplished this through the implementation of secondary models that predicted supplemental alloy properties for use as features in the final model to predict Clyne SCS. While artificial neural networks (ANNs) have historically stood out for their low complexity and high accuracy, we additionally evaluated the random forest (RF) and support vector regression (SVR) architectures with ANNs as alternative architectures to determine which model structure had the optimal prediction accuracy (16). Finally, we evaluated the accuracy of both models using cross-validation techniques with multiple folds of data. We demonstrated that while the final RF model with all features was less accurate than previous SCS-related research, the addition of the secondary predicted features; ductility, solidus, and GSF; improved the precision of our optimized SCS prediction model. These findings suggest a not-yet-explored relationship between SCS and these features which can be investigated in future research to produce potentially more accurate mathematical quantifications of SCS.

RESULTS

To develop a ML model to predict Clyne SCS values, we began with a dataset of alloy compositions, solidus. liquidus, freezing range, phases, surface energy, general stacking fault energy, ductility parameter, and Clyne SCS. Of these, we initially took alloy composition as the input to three models: a RF, ANN, and SVR.

Model Results

Before assessing, the output accuracy of the SCS models, hyperparameter tuning was completed to decide optimal hyperparameter values. For the RF model, we optimized three hyperparameters: number of estimators, maximum tree depth, and maximum number of features. We tested ten equally-spaced values for the number of estimators ranging from 2 to 1000 and found 667 estimators to be optimal. We tested six values between 2 and 30 for maximum tree depth (optimal value of 30), and nine values between 2 and 10 for maximum number of features (optimal value of 10). For the ANN model, we optimized the number of epochs and the learning rate through tuning. 560 epochs and a learning rate of 0.0001 was found to optimize output accuracy. Finally, we optimized the SVM's C value which we found to be 10.0.
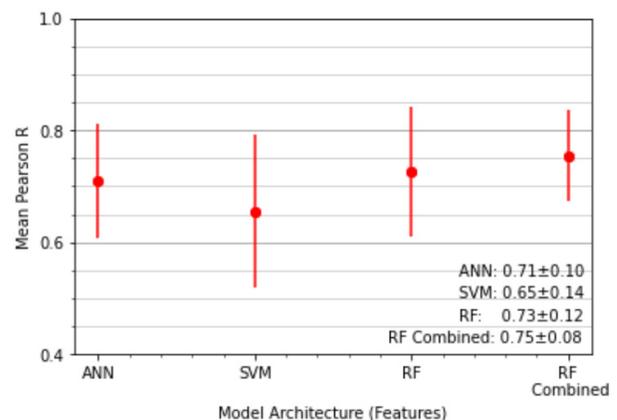

Figure 1: Mean Pearson R in 10-fold CV for each model architecture. The average Pearson R per fold of the ANN, SVM, RF, and combined RF pipeline showed that the RF with all inputs has the highest accuracies under the optimal combination of hyperparameters (**Table 1**).

Among the three model architectures, the RF model demonstrated the best performance, with a mean Pearson correlation coefficient-value of 0.73 ± 0.12. The ANN and SVR models delivered comparable performance, yielding mean Pearson correlation coefficient values of 0.71±0.10 and 0.65±0.14, respectively (**Figure 1**).

Addition of Secondary Attributes

Once the optimal architecture was determined, the additional information in the dataset was analyzed to further increase the accuracy our final prediction. To determine what attributes should be included in the input vector to our model, we generated a correlation matrix of the data to identify the relationship strength between the alloy composition and our desired output of SCS values. The matrix plotted the R^2 value between each feature and SCS, which measures the strength of a linear relationship between the features. The correlation values between Clyne SCS and the alloy composition information ranged between <0.01 and 0.17. Furthermore, the correlation coefficients between the other input variables, such as GSF and Ductility, and SCS were also relatively low, with the highest correlation coefficient being 0.22 (**Figure 2**). These low correlation values suggests that achieving high prediction accuracies for SCS may be challenging given the characteristics of the dataset. Nevertheless, we determined that including GSF, Ductility, and Solidus Temperatures as additional model parameters would provide the most useful information to SCS prediction models, as they exhibited the highest correlation coefficient magnitudes of 0.17, -0.15, and 0.22, respectively.

Based on the insights gained from the correlation matrix, the final model ensemble and training process were developed. We still focused using only the alloy composition as the initial input to minimize the need for extra alloy property simulations or computations. To achieve this, three additional ML models were designed to predict ductility, solidus, and GSF solely from the alloy compositions. The final SCS model
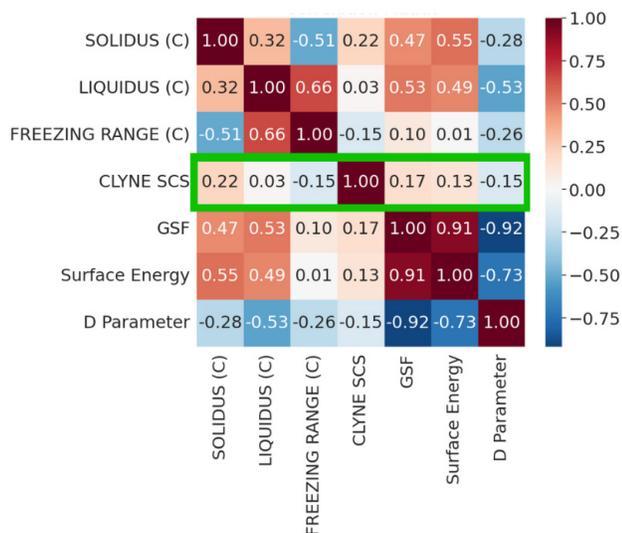

Figure 2: Feature correlation matrix. A correlation matrix for all possible additional inputs with highlighted values (in green box) for correlations corresponding to SCS. From this, we determined the most correlated features to include as additional inputs to the SCS model, specifically, solidus, ductility (D Parameter), and GSF.
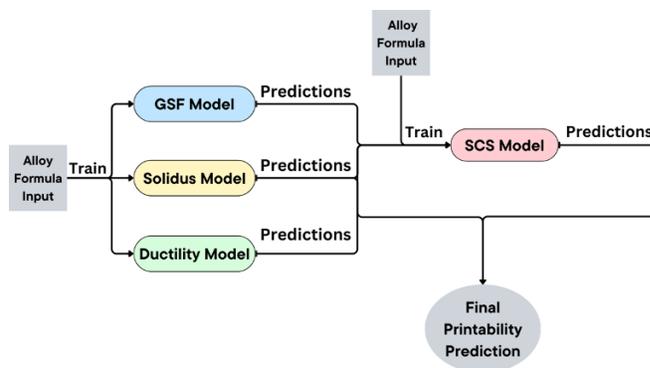

Figure 3: Final SCS model pipeline. Diagram depicting the training pipeline for all four models. We maintained the alloy formulas as the only inputs while creating a more informed feature vector for the SCS model. First, the formulas were used to train the secondary property models, which produced predictions for each property. Then the formula was aggregated with these properties to input a larger feature vector for the SCS model using the same 90-10 train-test split.

incorporated these predicted values along with the original alloy compositions when making its predictions (**Figure 3**). The ensemble, with models for the secondary properties feeding into our main SCS prediction model, enhanced the overall prediction performance and improve the model's effectiveness in assessing the printability and structural integrity of alloys.

Upon adding these model predictions to feed into our original random forest model, the primary RF model for SCS, predicted SCS values with a Pearson-R of 0.75±0.08. Comparing this to our previous structure that only handled alloy composition, the secondary property predictions decreased standard deviation of SCS value, indicating that the pipeline increased model precision on unseen test data.

Secondary Model Accuracy

Since the secondary RF and ANN models for ductility, solidus, and GSF reached extremely high accuracy (Pearson correlation coefficient-value of 0.99). Although such a high accuracy can indicate overfitting, overfitting did not appear to explain our levels of accuracy based on observing testing outputs. Firstly, 10-fold cross-validation was performed to evaluate the model's ability to predict unseen data. The mean Pearson correlation coefficient remained at 0.99, showing that overfitting was not likely in this situation. An example of overfitting in models that we compared to is a visible and significant decrease in the testing accuracy as the number of estimators continues to increase (**Figure 4**). As we did not see evidence of this, we concluded that our models converged at the accuracy of a 0.99 Pearson correlation coefficient-value.

Statistical Testing

The chosen accuracy metric, Pearson correlation coefficient directly compares the model output with the expected value and is useful in comparing model architectures. However, a statistical significance test is needed to assess whether the achieved Pearson correlation coefficient represents sufficient accuracy (21). In this work, we used t-tests to determine whether our predictions were significantly different than the true values for SCS. Given that
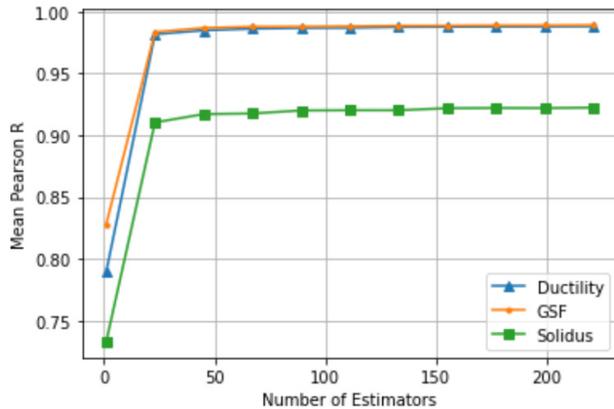
Figure 4: Mean Pearson R in 10-fold CV vs number of RF estimators. A graph of the average Pearson R versus the number of estimators showed constant accuracy, suggesting that the model is not overfitting.

the combined RF pipeline was assessed to have the highest accuracy, the t-test was performed on its predictions. The average p-value for 10 trials of model prediction and t-tests was 0.111, which is greater than a standard alpha threshold value of 0.05 (21). Thus, we failed to find a significance difference between predicted and actual values of SCS for our final combined RF pipeline.

DISCUSSION

Overall, this study aimed to predict the Clyne SCS of potential printing alloys based solely off their chemical composition. This can help manufacturers determine the strength of formulas without the need for empirical testing. In addition to this goal, we found introducing secondary feature inputs predicted through separated ML models, even those beyond the original equation, benefitted the ML predictions.

For our models, we determined an optimal way to quantify model error. A common method of quantifying error with numerical output is mean absolute error (MAE) which is described by the following equation (18) with n denoting the total number of data points the model predicts, y being the true SCS, and $\hat{y}$ being the predicted SCS.

$$MAE \ = \ \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad \text{(Equation 2)}$$

Unfortunately, MAE proved ineffective for our purposes because of the spread of our dataset. There does not appear to be a set range for SCS. In particular, values that differ from the mean by multiple magnitudes of standard deviation can bias error calculations, making it appear that we had extremely large errors relative to the mean. While these cases are important to consider, we opted to use an accuracy metric that accounted for their low occurrence rate, the Pearson correlation coefficient. Our switch was also motivated by other SCS prediction research that had success using Pearson correlation coefficient instead of MAE (12). Pearson's correlation coefficient is described by Equation 3 (12,21). In this equation, $\hat{y}$ denotes a predicted value, x denotes the mean predicted value, in the Pearson correlation coefficient, y denotes the expected value, and $\bar{y}$ denotes the mean expected value.

$$r \ = \ \frac{\Sigma(\hat{y}_i - x)(y_i - \bar{y})}{\sqrt{\Sigma(\hat{y}_i - x)^2 \Sigma(y_i - \bar{y})^2}} \qquad \text{(Equation 3)}$$

We found that the random forest model architecture performed more accurately than the ANN and SVM. This result may be due to the high dimensionality of the data, the existence of more outliers, or the low correlation or high complexity of relationships between the features and the SCS, all of which strengthen random forest models. The difference in average Pearson correlation coefficient between the random forest with all features and the one with only alloy compositions is statistically not significant. This lack of significance may be due to the small correlations in these added variables, and hence adding them into the model may not provide much more information to train the model. However, the lower standard deviation of the Random Forest with all features indicates that this model is more consistent in achieving this average. Moreover, the accuracy of the enhanced RF model was further tested by a statistical t-test which failed to find a difference between predicted and expected outputs by our final model.

Our model was limited by the information in our dataset. The dataset itself, although inclusive of alloys beyond Ti- and Ni- based alloys seen in past research, does not contain all possible elements that may form printable alloys. Thus, our model may not generalize to alloys containing elements outside of the simulated dataset. Expanding the included element set would further decrease the set of alloys that would need to be empirically tested. Furthermore, the data were limited by inaccuracies in the simulation software ThermoCalc itself, as the values generated may vary slightly from experimental reality. However, we note that it would be unfeasible to experimentally generate a dataset of this size.

Additionally, our study focused on Solidus temperature, General Stacking Fault Energy, and ductility as secondary inputs. Moreover, we further tested the achieved accuracy of the final RF model by a statistical t-test which failed to find a difference between predicted and expected outputs by our final model. In the future, other relevant or highly correlated secondary properties could be explored to further increase the accuracy of SCS prediction. Nonetheless, we believed these findings were still a meaningful to inform manufacturers with a final printability decision.

The final results of this experiment indicated several trends. Firstly, predictions were improved under a two-step pipeline in which other features are predicted and subsequently compounded with the initial composition data to calculate Clyne SCS. While prediction accuracy did not significantly increase, the relative precision of our predictions did increase, supporting our initial hypothesis that adding these features would aid SCS predictions.

Secondly, we found additional features used in the pipeline included features beyond those used through the preexisting Clyne SCS formula. Thus, the relationship between Clyne SCS and other physical properties such as GSF and Solidus can be further explored to describe the additional nuance supported by our findings on which features aid Clyne SCS predictions. Though our work displays how ML is poised to bypass the need for simulating these mechanical relationships,

future work may expand on our attempt to produce a more accurate mathematical relationship for SCS that factors measurements outside the Clyne SCS formula. For instance, the freezing range feature was equally correlated with SCS as ductility, although we only included the latter in our secondary model due to time constraints.

Finally, though SCS is usually reserved for cracking susceptibility predictions of materials, the methodology explored in this study can be extended to other material properties. For example, past research explains two other criteria for assessing high-temperature alloys interest: solid-state cracking and partitioning (4). While SCS predicts the material's affinity to develop weaknesses or faults during solidification, solid-state cracking determines the material's performance under high-stress loads (4). Partitioning describes potential molecular separations with certain alloy compositions that would lead to overall structural weakness and microstructure development in the final product. Since these additional criteria are closely related to the features handled in this paper and dataset, we argue that research following the same ML pipeline structure can be leveraged to develop predictive models for further quantification of these other weakness metrics.

MATERIALS AND METHODS
Models
We chose three models for our prediction of SCS: RF, ANN, and SVR. Additionally, all secondary models—ductility, solidus, and GSF—utilized an RF architecture. RF and ANN in particular are widely used ML algorithms adept at handling large datasets with several correlations. We aimed to compare the performance of these three models in predicting Clyne SCS.

The first architecture we implemented in our model was a random forest. An RF trains by drawing random bootstrap samples of data, which it uses to form an ensemble of decision trees based on patterns and relationships within the data. Splitting the data into random subsets limits the probability of unstable predictors and improves the accuracy of the model on previously unseen data (17). In a typical random forest, the model excludes unimportant features and selects the most important and correlated features. In this study, however, we assume that all component elements, relative to each other, may contribute substantially to the cracking susceptibility of a specific alloy. RFs have been utilized as a ML model for over twenty years, and its flexibility concerning input (numerical, ordinal, or nominal) and data distribution has encouraged frequent usage in many applications (18). RF is adept at handling large datasets such as ours efficiently and accurately due to its ensemble nature (19). The second architecture we chose was an ANN, which applies linear matrix multiplication as well as intermediate nonlinear activation functions to the input vector for each layer of the network. In this case, the rectified linear unit (ReLU) activation function was used for all layers, allowing the ANN to be capable of detecting complex nonlinear relationships. Although there has been extensive research into the advantages of DNNs versus SNNs, for our purposes, we have found that DNNs typically perform with significantly higher accuracy than SNNs (12). Therefore, we designed our Clyne SCS prediction model to use five hidden layers (10-30-50-100). Our model used the mean squared error loss function and the Adaptive Moment Estimation

(ADAM) optimizer, which allows for more precise gradient descent, set at an initial learning rate of 0.001. Finally, the third architecture we chose was SVR, which is commonly used to define acceptable errors in our model and determine an appropriate line, or the optimal hyperplane, to fit the data. The advantages of an SVR include being less prone to overfitting and outliers within the dataset (13). The kernel function that we used was the radial basis function (RBF), which determines non-linear regression and can determine multivariable relationships. The arguably most important hyperparameter choice for SVR is the value of Cost (C), which modifies the model's tolerance for points outside of the acceptable error margin. A greater C value reduces the risk of misclassification but also makes the model more likely to overfit our dataset. To find an optimal C value, we tested multiple candidates through a grid search and found the best solution, which gained the greatest Pearson correlation coefficient (R)-value.

Data Preparation
A dataset of compositions and important solidification parameters was generated to evaluate the relationship between cracking susceptibility and composition-dependent parameters. A Python script was used to generate >3,500 unique compositions in the refractory multi-principal element alloy space for thermodynamic evaluation. For each composition, a Scheil simulation was performed in ThermoCalc using the database TCHEA5 and a TC-Python script to automate the calculations (20). The shape of the fraction solid-temperature curve was fed into the Clyne solidification cracking susceptibility (SCS) predictive criterion, which could be compared across alloy compositions (11). The range of SCS for the dataset utilized in this study was approximately -1.6–22.8. The dataset contains essential information about alloy composition, including solidus temperature, liquidus temperature, freezing range, phases, surface energy, general stacking fault energy, ductility parameter, and Clyne SCS. The alloys studied in this dataset consist of the following elements: Zr, V, Ti, Nb, Mo, Ta, W, Hf, Ru, and Re. Notably, these elements possess relatively high melting points, making them valid potential replacements for traditionally machined high-temperature parts used in aerospace applications. It is important to mention that not all of the alloys in the dataset contain all of the aforementioned elements, and certain elements are only present in a limited number of data points. For this paper, the primary focus lies on solid-state and solidification cracking susceptibility as key factors in determining an alloy's printability. Consequently, the target variables for the predictive model are Clyne SCS and ductility parameter, with the latter being directly related to solid-state cracking susceptibility (11).

Given the objective of eliminating the requirement for manufacturers to obtain physical, simulated, or calculated data, our features would ideally only consist of the alloy composition. To preprocess these compositions, we split the data of alloy composition, originally represented as a string of elements and their percent mass, into separate input variables denoting the numeric percent composition for each element. To ensure consistent training, feature values were standardized by scaling to unit variance to address potential challenges due to differing unit scales.

Model Architecture	Hyperparameters: {Possible Values}	Optimal Combination
Random Forest	Number of Estimators: {2,112, 223, 334, 445, 556, 667, 778, 889, 1000}	667
	Maximum Tree Depth: {2, 7, 13, 18, 24, 30}	30
	Maximum Number of Features: {2, 3, 4, 5, 6, 7, 8, 9, 10}	10
ANN	Epochs: {$epochs \in Z$| $20 \leq epochs \leq 600$, $step = 20$}	560
	Learning Rate: {0.01, 0.001, 0.0001}	0.0001
SVM	C: [0.01, 0.1, 1, 10, 30, 60,100]	10.0

Table 1: Grid search CV combinations for each model architecture. The hyperparameter grid for the grid search for each model architecture, including the hyperparameter combination that was determined to optimize each architecture.

Error Quantification

Pearson correlation coefficient accuracy quantification, given by Equation 3, was done through 10-fold cross-validation (CV). This validation method splits the data into 10 folds following the usual randomized shuffling. The model is trained 10 times with a different fold becoming the test set and the remaining data as the training set. The model's final accuracy is then averaged across all folds. Unlike traditional holdout techniques, a 10-fold CV overcomes test-train split sensitivity by checking that the model is not overfitting to certain data splits. Overall, CV provides a more realistic measure of accuracy and ensures comprehensive coverage of the dataset trends during evaluation.

SCS Model Hyperparameter Tuning

To calculate the hyperparameters that optimized accuracy, a grid search was performed for all suggested model architectures using a 10-fold (90-10 train-test split) nested k-fold procedure. By iterating through a grid of possible combinations of hyperparameters, the optimal combination that would provide the highest test accuracy was determined and used in final model accuracy assessments (**Table 1**). Pearson correlation coefficient was recorded for each, and the best parameters were selected. The model was then retrained in a 10-fold cross-validation to record the final accuracy.

REFERENCES

1. Wang, C., et al. "Machine learning- in Additive Manufacturing: State-of-the-art and Perspectives." *Additive Manufacturing*, vol. 36, no.101538, Dec. 2020, https://doi.org/10.1016/j.addma.2020.101538.
2. Chowdhury, Sohini, et al. "Laser powder bed fusion: A state-of-the-art review of the technology, materials, properties & defects, and numerical modeling." *Journal of Materials Research and Technology*, vol. 20, 24 Aug. 2022, pp. 2109-2172, https://doi.org/10.1016/j.jmrt.2022.07.121.
3. Gupta, Nayanee, et al. "Additive manufacturing: status and opportunities." Science and Technology Policy Institute, Washington, Mar. 2012, cgsr.llnl.gov/content/assets/docs/IDA_AdditiveM3D_33012_Final.pdf.
4. Mullin, K. M., et al. "Rapid screening of single phase refractory alloys under laser melting conditions." *In Materials & Design*, vol. 238, 2024, p. 112726, https://doi.org/10.1016/j.matdes.2024.112726.
5. Kilkenny, Nancy S. "3D printed rocket launched using innovative NASA alloy." *NASA's Glenn Communication*. 2 Mar. 2023. www.nasa.gov/feature/glenn/2023/3d-printed-rocket-launched-using-innovative-nasa-alloy. Accessed 15 July 2023.
6. Oztan, Cagri, and Victoria Coverstone. "Utilization of additive manufacturing in hybrid rocket technology: A review." *Acta astronautica*, vol. 180, Mar. 2021, pp.130-140, www.sciencedirect.com/science/article/pii/S0094576520307001.
7. Leary, Martin. "Fundamentals of Laser Powder Bed Fusion of Metals", edited by Yadroitsev, Igor et al. Elsevier, Science Direct, 2021, pp. 597-620, https://doi.org/10.1016/B978-0-12-824090-8.00022-6.
8. Soysal, Tayfun, and Sindo Kou. "A simple test for assessing solidification cracking susceptibility and checking validity of susceptibility prediction." Acta Materialia, vol.143, 15 Jan. 2018, pp.181-197, https://doi.org/10.1016/j.actamat.2017.09.065.
9. Kou, Sindo. "A criterion for cracking during solidification." *Acta Materialia*, vol.88, 15 April 2015, pp. 366-374, https://doi.org/10.1016/j.actamat.2015.01.034.
10. Li, Z. M., et al. "Cuboidal γ'phase coherent precipitation-strengthened Cu–Ni–Al alloys with high softening temperature." *Acta Materialia*, vol. 203, no. 116458, 15 Jan. 2021, https://doi.org/10.1016/j.actamat.2020.10.076.
11. Clyne, T. W. et al. "The effect of melt composition on solidification cracking of steel, with particular reference to continuous casting." *Metallurgical transactions B*, vol.13, 1982, pp. 259-266, link.springer.com/article/10.1007/BF02664583.
12. Feng, Shuo, et al. "Using deep neural networks with small dataset to predict material defects." *Materials & Design*, vol. 162, 15 Jan. 2019, pp. 300-310, https://doi.org/10.1016/j.matdes.2018.11.060.
13. Noble, William S. "What is a support vector machine?" *Nature Biotechnology*, vol. 24, no. 12, Dec. 2006, pp.1565-1567, https://doi.org/10.1038/nbt1206-1565.
14. Liu, Jikai, et al. "Current and future trends in topology optimization for additive manufacturing." *Structural and multidisciplinary optimization*, vol. 57, 3 May 2018, pp. 2457-2483, https://doi.org/10.1007/s00158-018-1994-3.
15. Kotadia, H. R, et al. "A review of laser powder bed fusion additive manufacturing of aluminum alloys: Microstructure and properties." *Additive Manufacturing*, vol. 46, no. 102155 Oct. 2021, https://doi.org/10.1016/j.addma.2021.102155.
16. Abiodun, O. I. et al. "State-of-the-art in Artificial Neural Network Applications: A survey." *Heliyon*, vol. 4, no. 11, 2018, https://doi.org/10.1016/j.heliyon.2018.e00938.
17. Segal, Mark R. "Machine Learning Benchmarks

and Random Forest Regression." *UCSF: Center for Bioinformatics and Molecular Biostatistics*, 2004, escholarship.org/uc/item/35x3v9t4.

18. Smith, Paul F. et al. "A comparison of random forest regression and multiple linear regression for prediction in neuroscience." *Journal of neuroscience methods*, vol. 220, no. 1, 30 Oct. 2013, pp. 85-91, https://doi.org/10.1016/j.jneumeth.2013.08.024.

19. Breiman, Leo. "Random forests." *Machine learning*, vol. 45, Oct. 2001 pp. 5-32, https://doi.org/10.1023/A:1010933404324.

20. Thermo-Calc Software TCHEA High Entropy Alloys Database version 5, thermocalc.com/products/databases/high-entropy-alloys/. Accessed 5 August 2023.

21. Obilor, Esezi Issac and Eric Chikweru Amadi. "Test for Significance of Pearson's Correlation Coefficient (r)", *International Journal of Innovative Mathematics, Statistics & Energy Policies*, vol. 6, 2018, pp. 11-23, seahipaj.org/journals-ci/mar-2018/IJIMSEP/full/IJIMSEP-M-2-2018.pdf.