International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 1 ǁ 2022 ǁ PP. 37-43

Using a machine learning model to predict the effect of parameters on die casting quality

Dr. Shueiwan H. Juang[1];David A. Kafando[2]

[1]Professor; [2]graduate student
National Taiwan Ocean University, Department of Mechanical and Mechatronic Engineering, Taiwan (ROC)
Corresponding author: shjuang@mail.ntou.edu.tw

ABSTRACT
In the present work machine learning is used to build a regression artificial neural network model, which predict the porosity, based on the casting weight. Then, the machine-displayed parameters and the setting parameters are linked using polynomial regression. In addition, a classification model is built using artificial neural network to predict the quality of the casting based on the displayed parameters. So far, this research is only a preliminary machine-learning model and it is still unable to achieve self-learning. In total more than nine hundreds datasets are involved in this study, these datasets have been collected on field at a die casting company. Using the research results obtained from three models it would be easier to adjust the machine parameters to get a low-porosity casting. The first model (Model A) was built successfully with a mean absolute error that reaches 0.06. The second model (Model I) was evaluated using R squared and it reaches 0.63 to 0.99 for the different outputs (machine displayed parameters). The third model (Model II) has an accuracy of 80%. These models when combined could simplify the production line, boost the yield quality, hence increase the productivity.

Keywords: *Die casting, Machine learning, ANN, Casting quality, porosity*

Date of Submission: 15-01-2022 Date of acceptance: 30-01-2022

I. INTRODUCTION

The service revenue of Artificial intelligence (AI) has increased by 13% in 2020 and is expected to increase by 17.4% in 2021 [**1**]. This intelligence is being incorporated into the casting industry to boost the production, also to simplify the production and the casting testing.Up until now, the die-casting industry relies on the experience of die-casting production masters and traditional trial and error methods to adjust process parameters improving the quality of castings. Many researchers have worked to improve the casting industry using the artificial neural network. Rai et al [**2**] used artificial neural network to develop a model with an accuracy of 96.5% have proven that the melt temperature, the mold temperature and 1st phase velocity influence the casting porosity, the filling time and the solidification time. Jeongsu Lee et al built an artificial neural network model, which predict the quality of casting with 47 inputs variables. Their model could reach a 96.9% overall accuracy, and 99.875% accuracy for the production success cases (799/800).The summary of the results is shown in **Figure 1**[**3**].

(*a*)

	Predicted Production success	**Predicted** Production failure
Actual Production success	799	1
Actual Production failure	30	170

Figure 1: Confusion matrix of the prediction model.

This work used regression artificial neural network to establish a prediction model of the porosity using the weight as input, in which the relationship between the setting parameters and the displayed parameters is established using polynomial regression. Finally, a classification neural network model is builtto predict the quality of the casting based on the response parameters of the machine.

II. METHODOLOGY

The aim of this work is to use the programming Language-Python 3 constructing AI models with machine learning to predict the impact of die-casting process parameter variation on the quality of die-casting parts, and to monitor die-casting production information, improve die-casting yield and increase the production rate. The flowchart used in the study is shown in**Figure 2**.

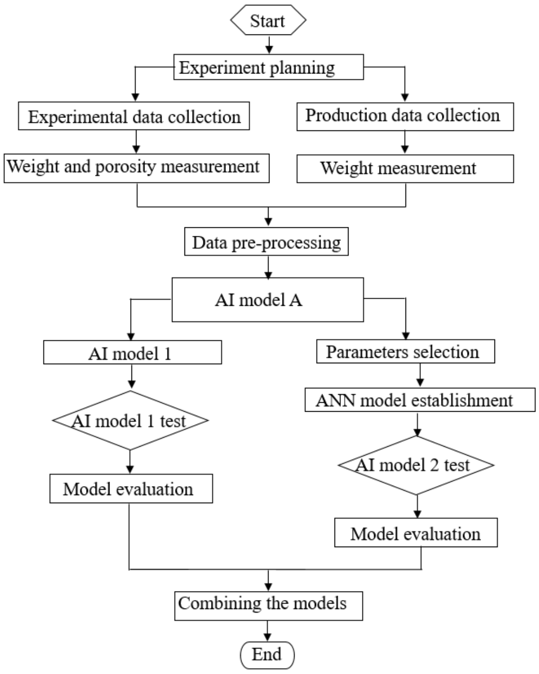

Figure 2: Research flow chart.

2.1 DATA COLLECTION

The data is collected from a die caster in Taiwan. The casting used in this study is the housing of a pneumatic tool that weights around 410 g (gating and venting systems not included) as shown in**Figure3**. There are in total of 866 datasets collected from the mass production. In addition, the setting parameters such as high velocity starting position, hydraulic pressure, slow velocity phase and biscuit thickness were adjusted in order to study their influences on the casting weight and the porosity. The experiment design is listed in**Table 1**. There are 40 castings collected and have undergone a measurement of porosity. The die casting alloy usedin the study is aluminum A383 (ADC 12).

Figure 3: Housing of pneumatic tool used in the study.

Table 1: Experiment design of die casting.

Experiment index	Parameters	Values	No of shots
A	High velocity position (mm)	270	05
B	High velocity position (mm)	240	05
C	Hydraulic pressure (kgf/cm^2)	130	05
D	Hydraulic pressure (kgf/cm^2)	100	05
E	Slow velocity phase (m/s)	0.315/0.75/1.0	05
F	Slow velocity phase (m/s)	0.25/0.6/0.8	05
G	Biscuit thickness (mm)	25	05
H	Biscuit thickness (mm)	10	05

2.2 SELECTION OF FEATURE PARAMETERS

Usually in data science, it needs a way to build a model which includes only the most important features from hundreds or even millions of features. There exist three main benefits. First, it makes the model easier to interpret by reducing the number of variables in the analysis. Second, the variance of the model can be reduced, thereby reducing overfitting. Finally, the computational cost (and time) of training the model can be reduced [**4**].In this work, there are 6 setting parameters, such as cylinder pressure, high speed, high speed switching point, intensification start position, injection delay, and biscuit thickness, and 28 displayed parameters provided by company. The goal of selecting characteristic parameters is to select key parameters that affect the quality of castings. Reference literature review and industry commonly used parameters to select the variables of the artificial intelligence analysis model.

Correlation coefficient

The correlation coefficient is a number that indicates the strength of the relationship between two variables. There are many types of correlation coefficients, the most common of which is the Pearson coefficient represented by the Greek letter ρ. It is defined as the covariance between two variables divided by the product of the standard deviations of the two variables, as shown in Equation (1). The correlation matrix is a graph showing the relationship between many variables. **Figure 4** shows the correlation matrix of the data used in this work.

$$\rho(X,Y) = \frac{\mathrm{COV}\ (X,Y)}{\sigma_X \sigma_Y} \qquad (2)$$

in which $\rho(X,Y)$denotes the Pearson coefficient; COV (X, Y) is the covariance between X and Y; and σ_X and σ_Y are the standard deviations of X and Y, respectively.

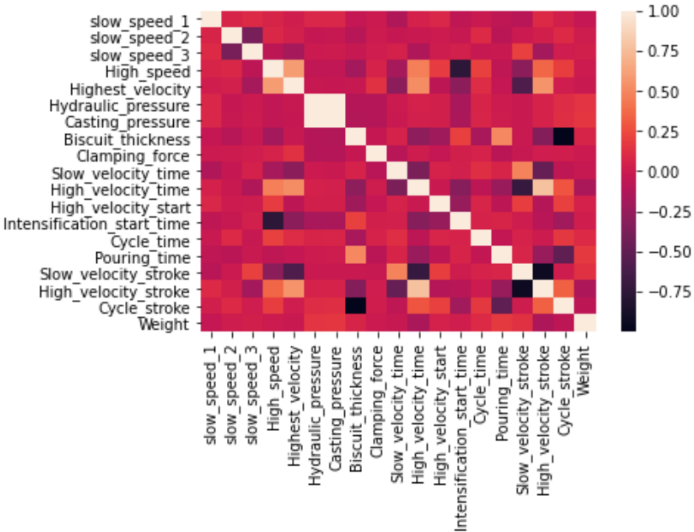

Figure 4: Correlation matrix.

According to the correlation matrix, by considering the correlation coefficient of the displayed parameters with a weighting value greater than 0.15 and lower than -0.15, the parameters such as high speed, cylinder pressure, casting pressure, biscuit thickness, pouring time, overall stroke, intensification switching time, high velocity stroke, high velocity switching time, and cycle time are selected as feature parameters. Those parameters have a certain relationship with the casting weight.

III. RESULTS AND DISCUSSION

3.1 MODEL A

Model A is established by using regression artificial neural network. Its purpose is to allow die casters to estimate the porosity of castings based on weight. Therefore, the input of the network is the weight of the casting, and the output is the porosity.**Figure 5** showed the relationship between weight and porosity. It can be obviously observed that the porosity variation is like the opposite of the weight. The correlation coefficient between weight and porosity of castings is -0.885. Therefore, there is a strong correlation between the two. The distribution of variables shown in **Figure 6** revealed a clear linear correlation between the data.The network construction of this model is very simple. It used an input neuron, a hidden layer containing 40 neurons, and an output layer. This model used the mean absolute error (MAE) and the mean square error (MSE) for evaluation.

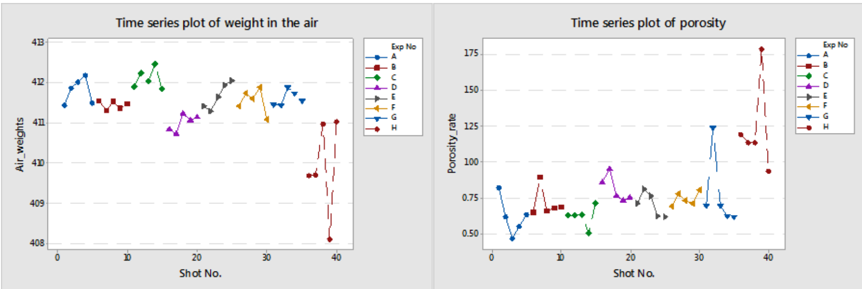

Figure 5: Data distribution of weight vs. porosity.

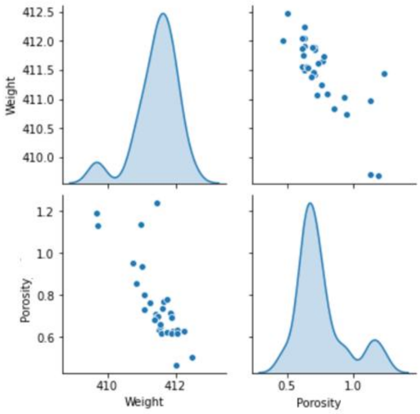

Figure 6: The distribution of the variables.

Figure 7 showed the MSE and MAE of the training dataset and the validation dataset along with epochs. The epochswas initially set to 500. The MAE reached 0.0430 and MSE reached 0.0035 at the last epoch. The low values of MAE and MSE indicated that the model may imply overfitting, which means that the model will not be able to accurately predict the porosity of new castings. Therefore, by applying the early stopping method, the model stops training when it converges. **Figure 8** showed that the model stopped training in the 319th epoch, and its MAE and MSE values were 0.0611 and 0.0075, respectively. This makes the prediction of the model more reliable and performs better in predicting the porosity of new castings.

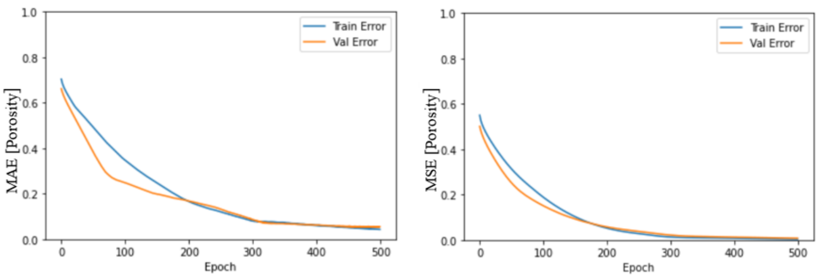

Figure 7: MAE and MSEaccording to the epoch.

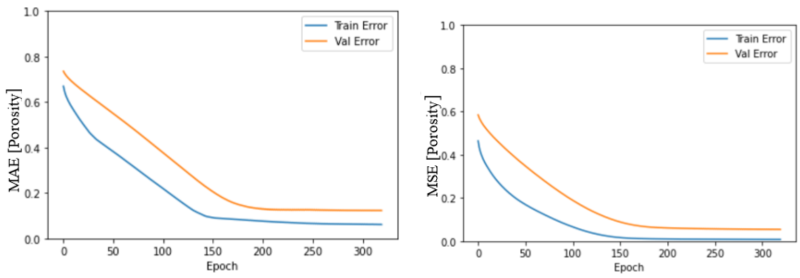

Figure 8: MAE and MSE according to epochs.

3.2 MODEL I

Model I is established by using polynomial regression analysis to find the relationship between the setting parameters and the machine response parameters. Equation (3) is the formula of polynomial regression analysis, where x_1~x_6 respectively indicate the setting parameters of cylinder pressure, high speed, high speed switching position, intensification start position, injection delay, biscuit thickness, etc.; y_i is the machine response parameter, which is selected as the output. This study uses polynomial regression analysis to find the relationship between the two types of die casting parameters, and uses R-squared (R^2) as the model evaluation coefficient. R^2 is a statistical measure that represents the proportion of variance of a dependent variable explained by one or more independent variables in the regression model, as shown in equation (4). The larger the value, the better the fit of the regression model.

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \cdots + \beta_{11} x_6 + \beta_{12} x_6^2 \ (i = 1, 2, \cdots, 8) \qquad (2)$$

where$i$ represents the number of the displayed parameter.

$$R^2 = 1 - \frac{\Sigma(Y_{actual} - Y_{predict})^2}{\Sigma(Y_{actual} - Y_{mean})^2} \qquad (3)$$

in which Y_{actual} is the actual value of the displayed parameter, $Y_{predict}$ is the predicted value andY_{mean} is the mean of the Y values.

Table 2 pointed out the evaluation results of Model I. The R-squared values of different outputs are all greater than 0.600, indicating that the prediction results of this model are reliable, while the R-squared of the casting pressure and cycle stroke is very high, both being 0.999.

Table 2: R^2 value of Normalized variables.

	R^2
High speed	0.872
Hydraulic pressure	0.996
Casting pressure	0.999
Biscuit thickness	0.635
High velocity stroke	0.999
Cycle stroke	0.999
High velocity time	0.745
High velocity switch time	0.811
Pouring time	0.890
Int. switch time	0.925

3.3 MODEL II

Model II is established by using the artificial neural network classification. It is a model for predicting the quality of die-casting parts. It aims to predict the quality of die-casting parts based on the parameters displayed by the machine. Through the prediction of Model A, the quality of the casting is obtained. Castings with porosity greater than 0.915% are considered NG (unqualified) and labeled with the number 0; castings less than or equal to 0.915% are considered good and labeled with the number 1. The artificial neural network classification model included an input layer, where 11 neurons represent the number of input variables; three hidden layers were used, and each layer contains 15 neurons; and the hidden layers were separated from each other by dropout, and the neurons are randomly discarded The connection forced the network to find new paths and generalize them. The rectified linear unit (ReLU) is the activation function used in the input and hidden layers.**Figure 9** showed a clear representation of the neural network.

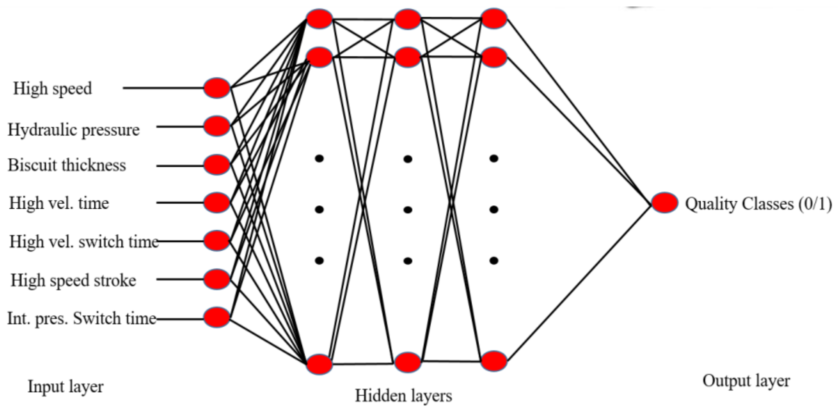

Figure 9: Classification neural network.

Table 4 indicated the prediction results of the training dataset and the test dataset. The prediction results of model II can reach an accuracy of 80%. Therefore, the quality of the input is better in terms of the impact on the quality of the casting. The confusion matrix of the prediction results was shown in **Figure 10**. It proved that this model not only has a higher accuracy but is also able to classify the NG casting. The recall is 93.81%.

Table 4: Model evaluation summary.

Accuracy	0.8000 (80.00%)
Precision	0.8347 (83.47%)
Recall	0.9381 (93.81%)
AUC	0.5556 (55.56%)

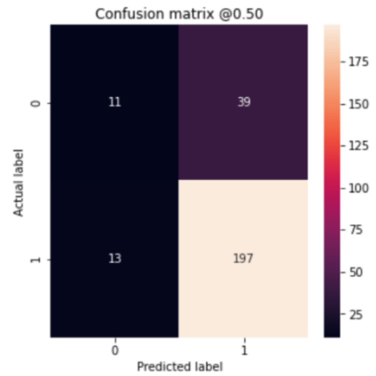

Figure 10: Confusion matrix.

The weight is measured very easily. An automatic balance can be added to the production line which could measure the weight as the production goes on. Those data can be connected to a computer to monitor the porosity using model A prediction. In the same perspective, the machine displayed data can also be used to predict the quality of the part using Model II. In case the quality is not as expected, the polynomial regression model make it easy to tune the setting parameters using their output function, and referring to the regression coefficient.

IV. CONCLUSION

- This study successfully established a regression neural network model, which predicts the porosity of castings based on density. The neural network model is very simple, using only one hidden layer and few neurons. The model used the average absolute error for evaluation, and its value reached 0.06. The research results showed that there is a high correlation between the weight of the casting and the porosity.
- This research successfully realized the predictive analysis of the second-order polynomial regression model, predicting the display parameters of the machine according to the setting parameters of the machine, and it allows understanding the setting parameters that have the greatest influence on the display parameters according to the regression coefficients. The study results showed that the R^2 of all outputs is higher than 0.6, among which the R^2 value of casting pressure and cycle stroke is as high as 0.999. Those results also

pointed out that the high plunger speed affects the many display parameters of machines, such as hydraulic pressure, casting pressure, biscuit thickness, high speed, high speed switching time, high speed time, high speed stroke, intensification pressure switching time, etc.

- The binary classification neural network model constructed in this study used the porosity generated by the regression neural network as the output. Importing in-depth analysis and hyperparameter tuning helps to optimize the model and make it more reliable. This study pointed out that the quality of input dataset would affect the accuracy of the model. Using high-speed, hydraulic pressure, biscuit thickness, high-speed time, high-speed switching time, intensification pressure switching time, as input parameters (display parameters), the accuracy can reach 80%, and the output was classified in binary, that is, 0 = NG and 1 = good.
- In order to use weight as an indicator of quality, it is necessary to ensure that the trimming of the runner and overflow of each shot should be consistent to avoid excessive variation in the measured weight data.

REFERENCES

[1]. www.idc.com, "IDC Forecasts Improved Growth for Global AI Market in 2021" [Online]

[2]. Rai J. K., Lajimi A. M., and Xirouchakis P, 2008, "An intelligent system for predicting HPDC process variables ininteractive environment", Journal of Materials Processing Technology, pp.72-79.

[3]. Lee J., Lee Y.C., and Kim J.T., 2021, "Migration from the traditional to the smart factory in the die-casting industry: Novelprocess data acquisition and fault detection based on artificial neural network", www.elsevier.com

[4]. Karkare R. S., "machine learning approaches to manufacturing and materials: applications to semi-supervised,unbalanced and heterogeneous data problems", 2019, thesis.