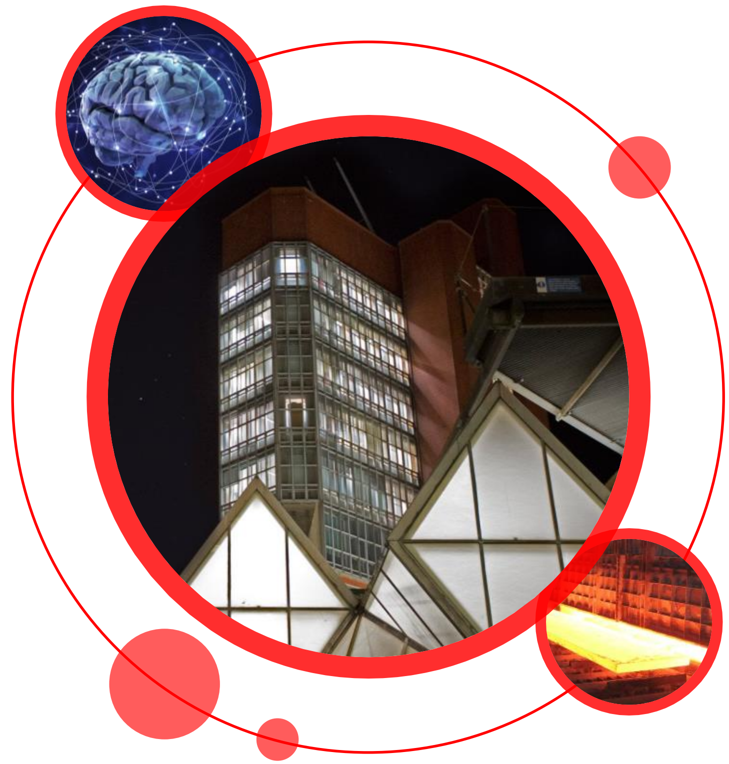
Prediction of Mechanical Properties of Low-Carbon Hot Rolled Steel Plates Using Machine Learning Methods

Xiaoan Yang, Jenny Shepherd, Qing Tao, Hongbiao Dong

School of Engineering, University of Leicester, LE1 7RH,
United Kingdom

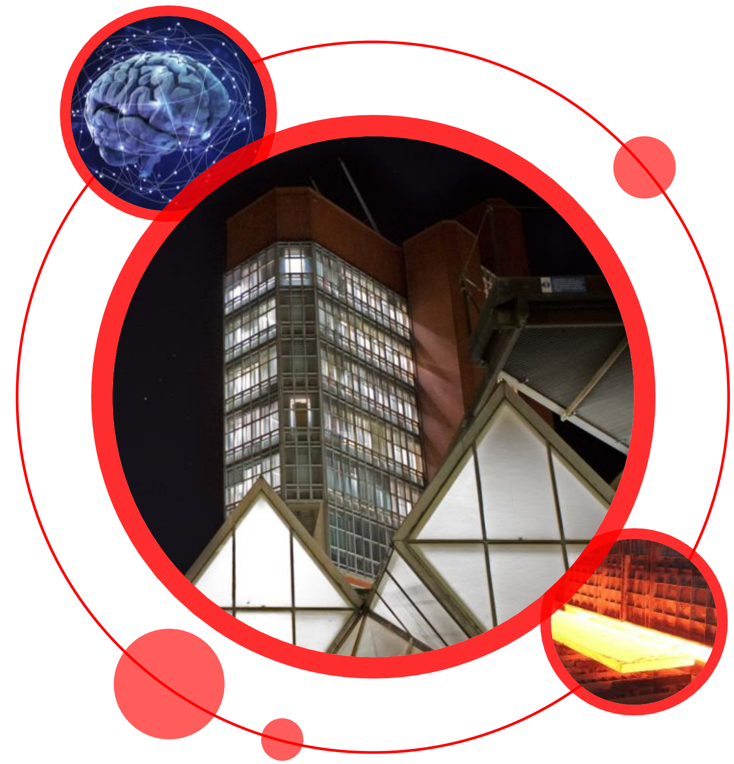
Content

1. Background
2. Prediction of yield to tensile ratio
3. Determination of key influential factors
4. Conclusion

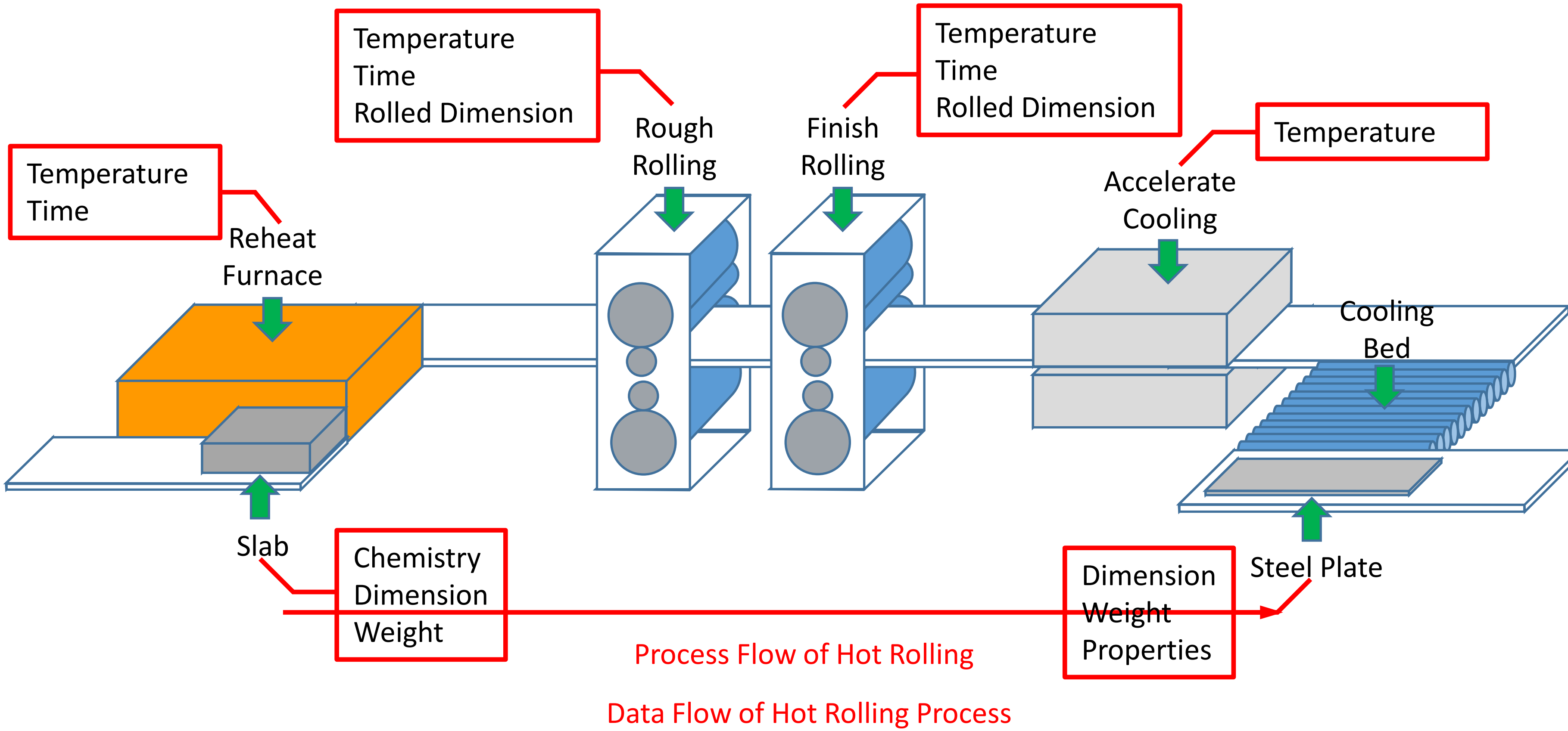


Content

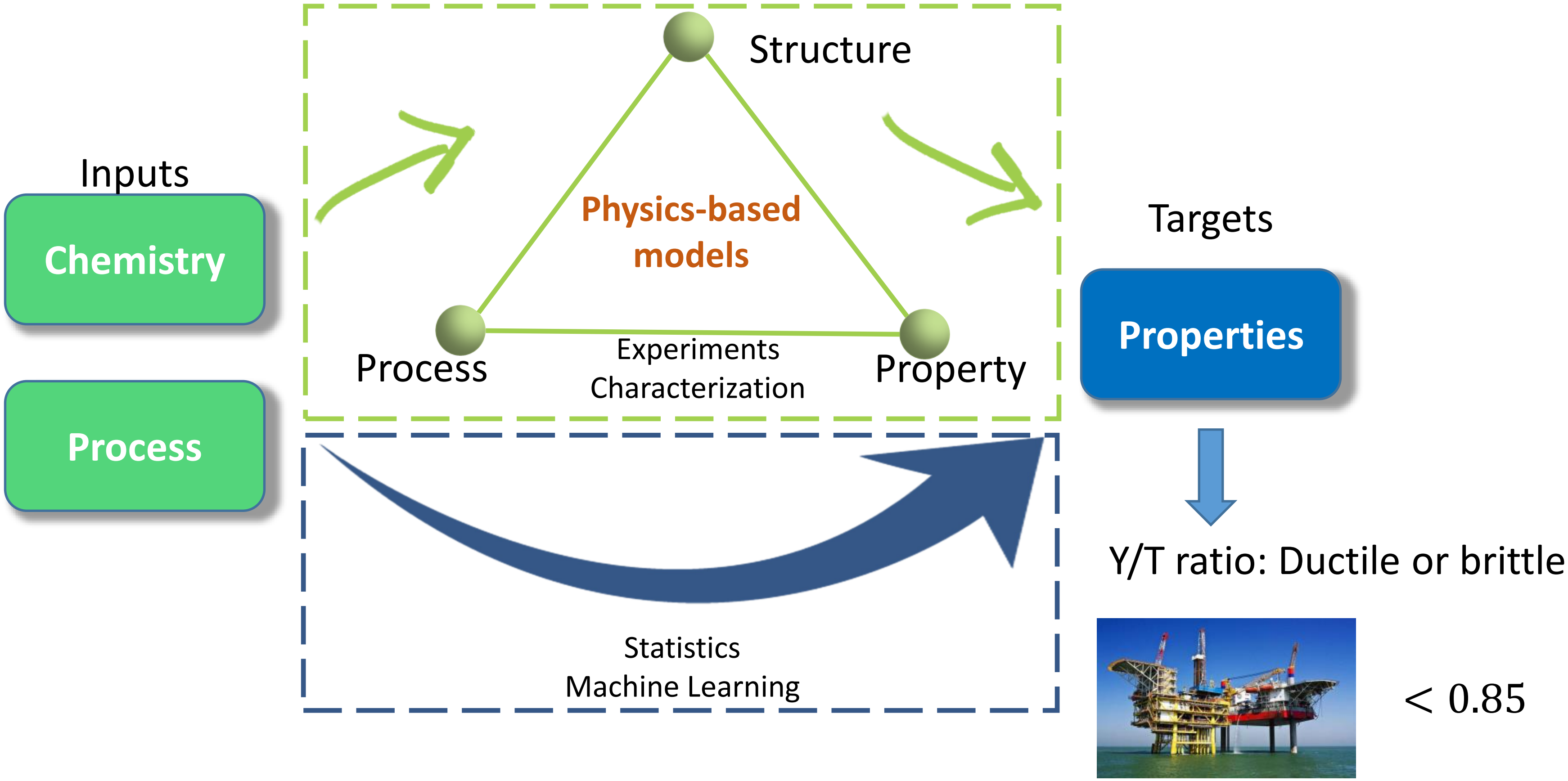
1. Background
2. Prediction of yield to tensile ratio
3. Determination of key influential factors
4. Conclusion



Process flow and data flow of hot rolling



Industrial Need and Aim for This Study



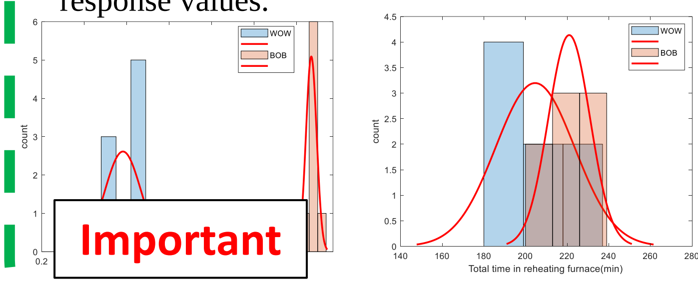
Benefits:

- Reduction in the cost of characterization of steel
- Efficient alloy design and process optimization

Introduction to iGate - Statistic Method

initial Guided Analytics for parameter Testing and control band Extraction (iGATE) gives an indication of the importance of one variable in predicting another

Null-hypothesis(H_0):
Input variable X that produced the highest response values have the same distribution as those values of X that produced the lowest response values.



Reject

Accept

8 groups of data with highest values of response variable (BOB)
8 groups of data with lowest values of response variable (WOW)

Sort the input variable X and corresponding output Y in ascending order of Y.

$(X_1, X_2 \dots, X_{16})$

$(Y_1, Y_2 \dots, Y_{16})$

Sort the input variable X and corresponding output Y in ascending order of X.

$\left(X_{(1)}, X_{(2)}, \dots, X_{(16)}\right)$

$\left(Y_{(1)}, Y_{(2)}, \dots, Y_{(16)}\right)$

Tukey-Duckworth test

$$s = \begin{cases} 0, & \text{if } Y_{(1)} = Y_{(16)} \\ s_l + s_u, & \text{otherwise} \end{cases}$$

Total number of outputs equal to $Y_{(1)}$ and $Y_{(16)}$

Wilcoxon rank sum test

$$w_u = \sum_{i=1}^{8} R_i \,,\; w_l = \sum_{i=9}^{16} R_i$$

The rank of X_i in the ascending order of X

Linear regression

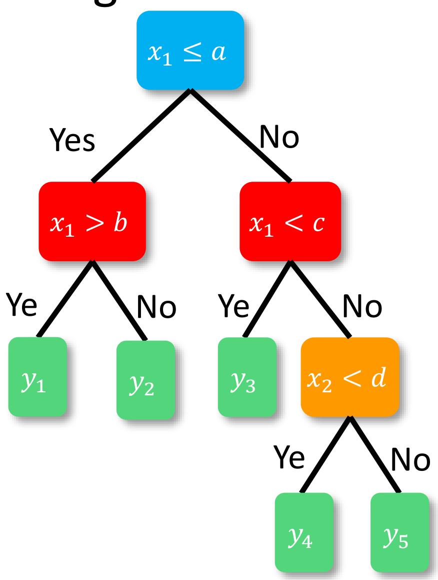
R^2

Stein S, Leng C, Thornton S, Randrianandrasana M. A guided analytics tool for feature selection in steel manufacturing with an application to blast furnace top gas efficiency. Computational Materials Science. 2021 Jan 1;186:110053.

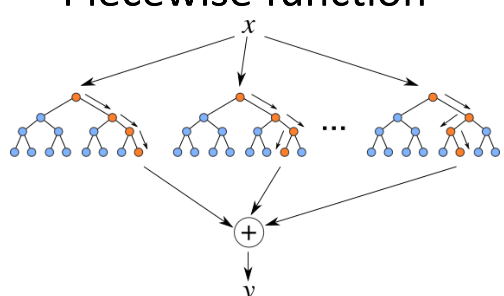
Introduction of Artificial Decision Tree and Neural Networks

"function approximation"

Regression Tree

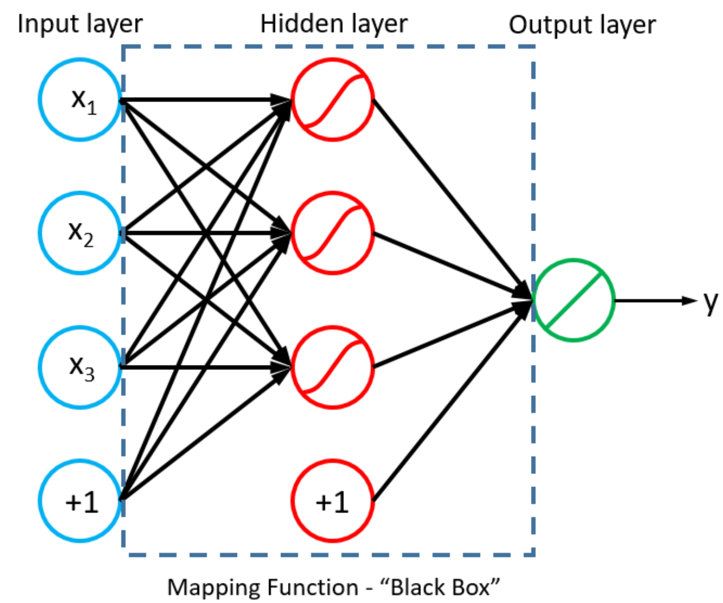


Piecewise function



Multiple trees = Random Forest

Artificial Neural Networks



$$h_k = \varphi\left(\sum_{j=1}^{m} w_{kj} x_j + b_k\right)$$

Continuous function

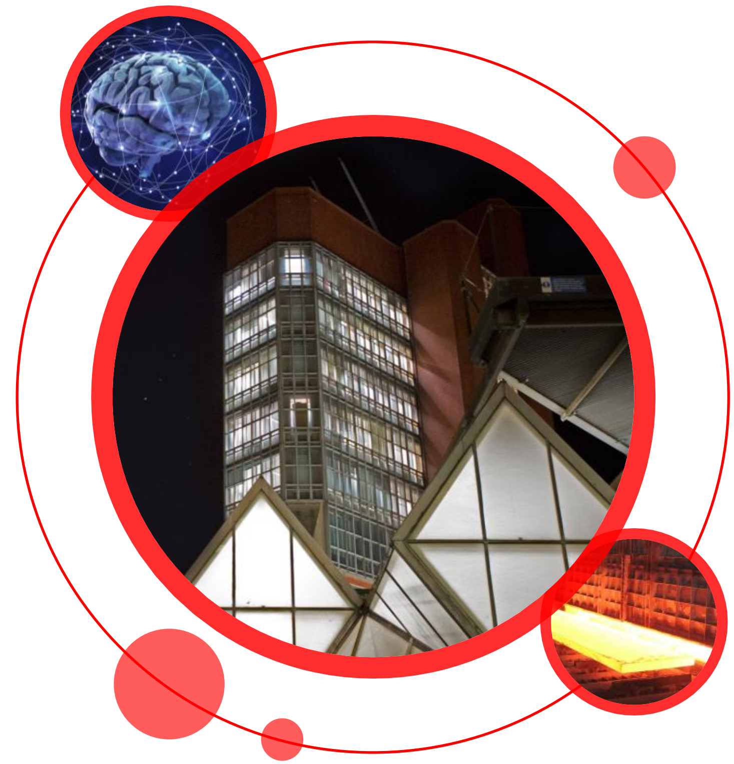
Content

1. Background

2. Prediction of yield to tensile ratio

3. Determination of key influential factors

4. Conclusion



Data overview

No.	Variable	No.	Variable
1	Slab thickness(mm)	24	Finishing rolled width(mm)
2	Slab width(mm)	25	Finishing rolled length(mm)
3	Slab length(mm)	26	Accelerate cooling T(°C)
4	Slab weight(t)	27	Thickness of rolled plate(mm)
5	Reheating T(°C)	28	Width of rolled plate(mm)
6	Heating zone T(°C)	29	Length of rolled plate(mm)
7	Total time in heating section(min)	30	Weight of rolled plate(t)
8	Soaking zone T(°C)	31	C(wt%)
9	Total time in soaking zone(min)	32	Mn(wt%)
10	Total time in reheating furnace(min)	33	P(wt%)
11	Rough descaling P(Mpa)	34	S(wt%)
12	Finish descaling P(Mpa)	35	Si(wt%)
13	Rough rolling T(°C)	36	Cr(wt%)
14	Rough descaling frequency	37	Cu(wt%)
15	Number of rough rolling pass	38	Ni(wt%)
16	Rough rolled thickness(mm)	39	Nb(wt%)
17	Rough rolled width(mm)	40	Mo(wt%)
18	Rough rolled length(mm)	41	V(wt%)
19	Inlet T in finishing mill(°C)	42	Ti(wt%)
20	Finish rolling T(°C)	43	B(wt%)
21	Finshing descaling frequency	44	Al(wt%)
22	Number of finishing rolling pass	45	Ca(wt%)
23	Finishing rolled thickness(mm)	46	As(wt%)

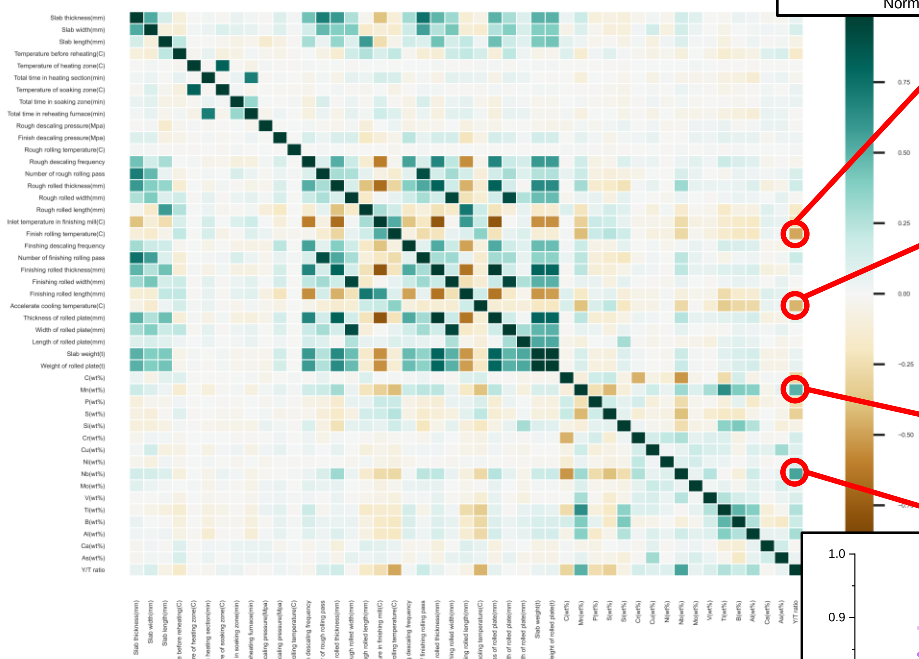
12000 + Industrial data with 46 input variables

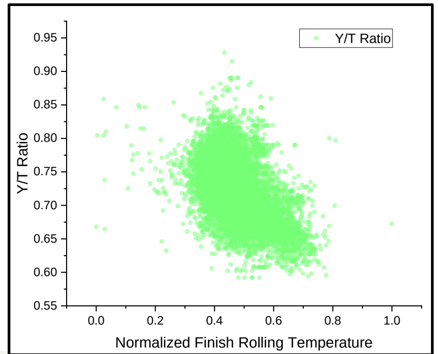
- Slab dimension
- Reheating parameters
- Rolling Parameters
- Plate dimension
- Chemistry

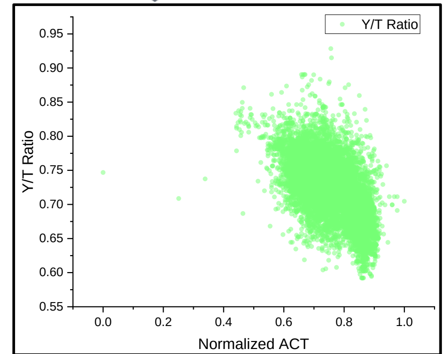
Data pre-treatment

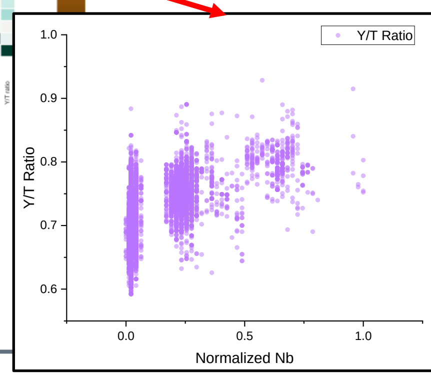
Outliers Removal(Tukey Fences)

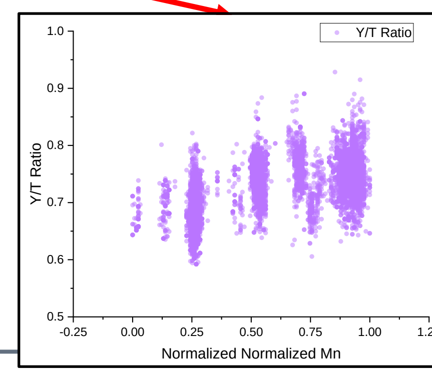
Linear correlation check





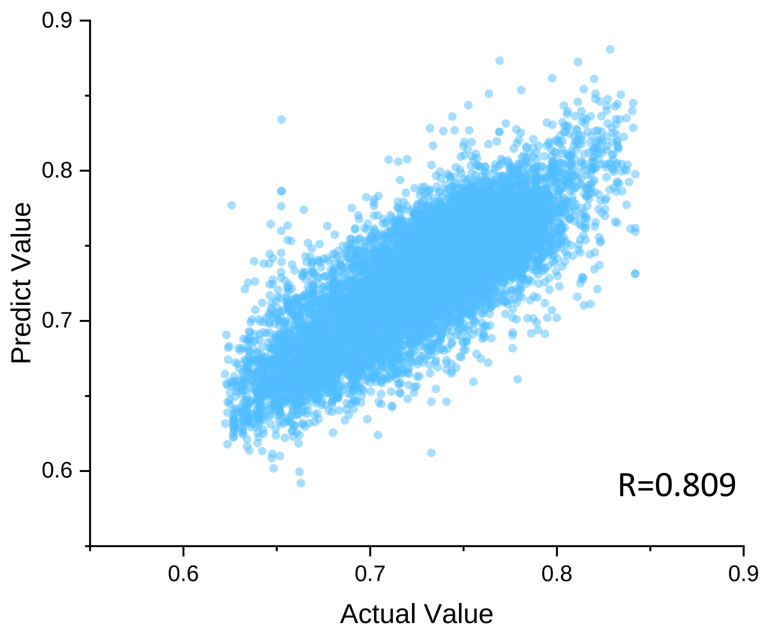




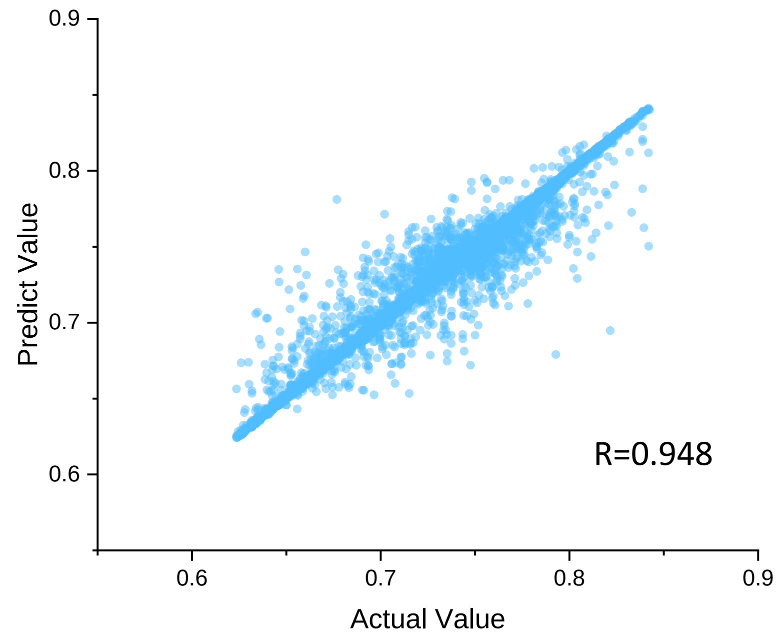


Normalization(Min-Max): $z_i = \frac{2(x_i - x_{\min})}{x_{max} - x_{min}} - 1$

Comparison of prediction accuracy: Neural Networks *vs* Regression Trees

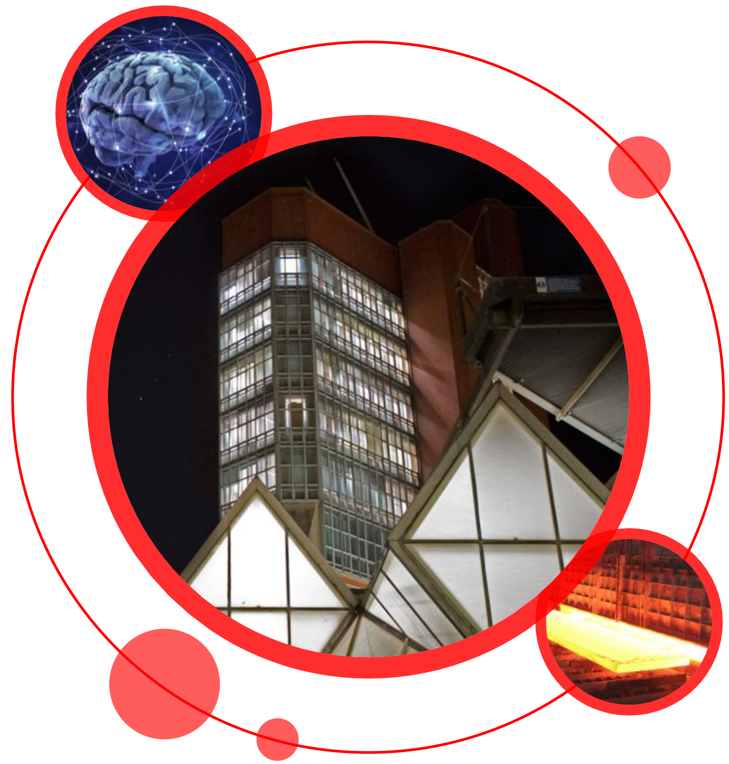


Neural Network



Regression tree
Extreme gradient boosting(Xgboost)

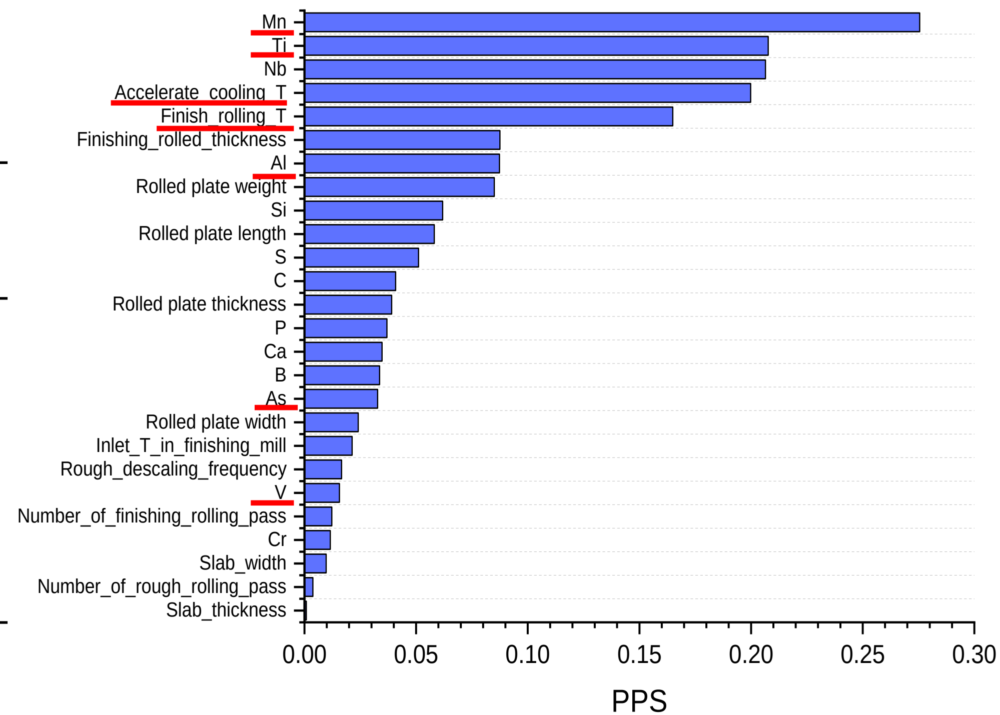
Content



Key influential features ranking using iGATE and PPS

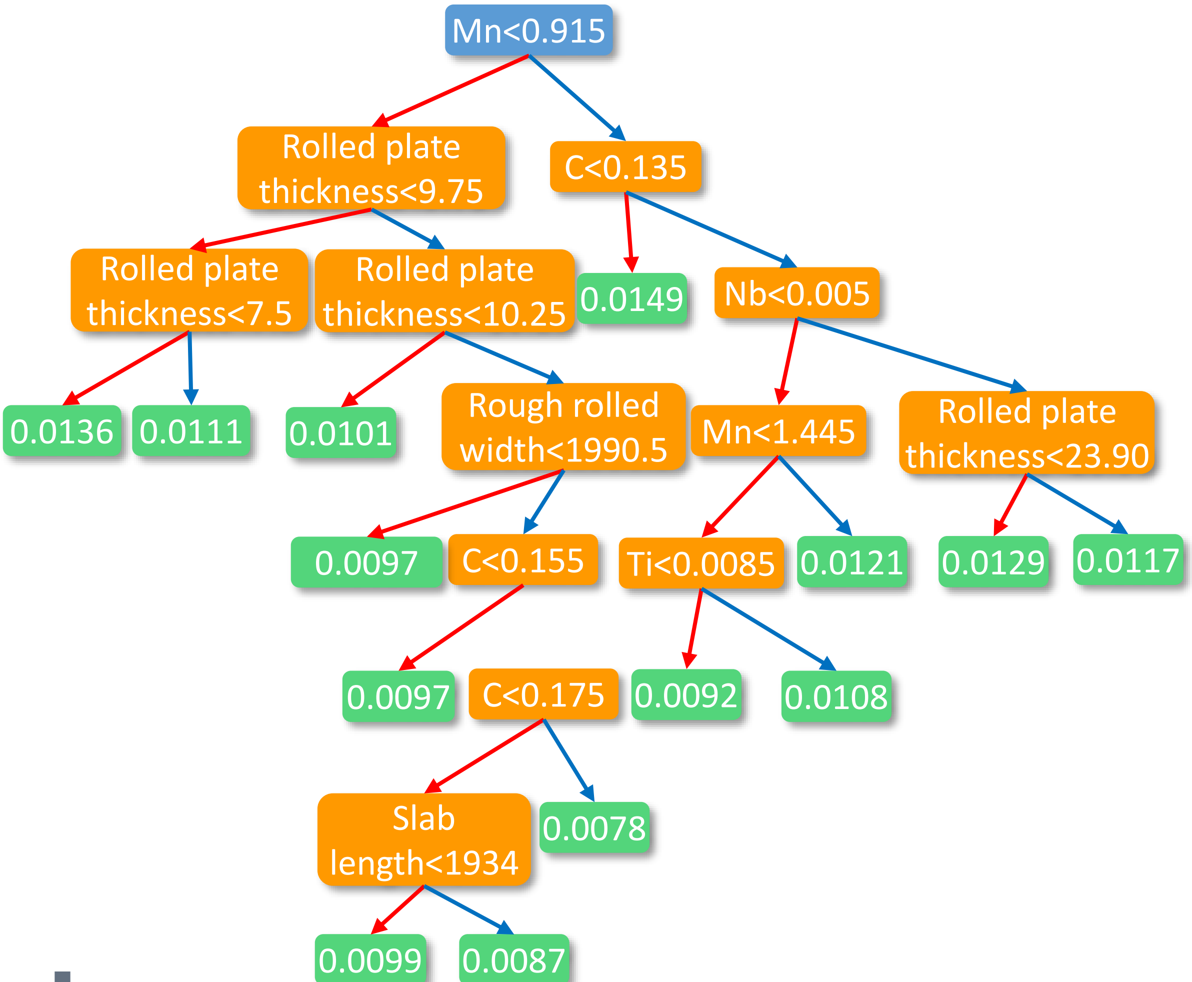
Features selected by iGATE

Feature	Count Tukey Duckworth	p.Values Wilcoxn Sum	R square
Accelerate Cooling T	10	0.005117	0.314
Mn	16	0.0008146	0.256
Finish Rolling T	6.5	0.0267439	0.159
Al	9.5	0.0079315	0.077
Ti	8.5	0.0074693	0.068
V	11	0.0015077	0.048
As	9	0.0090091	0.008

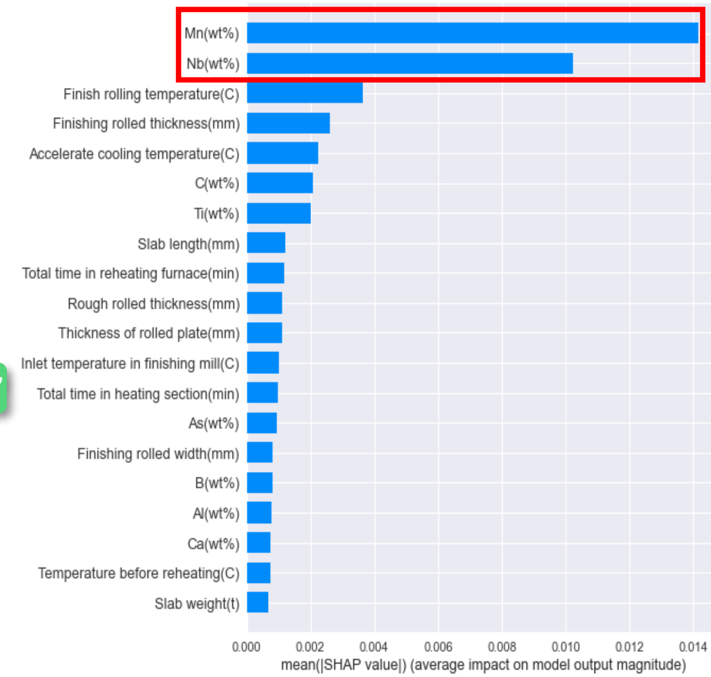


Feature importance ranking by Predictive power score(PPS)

Key influential features ranking based on Xgboost

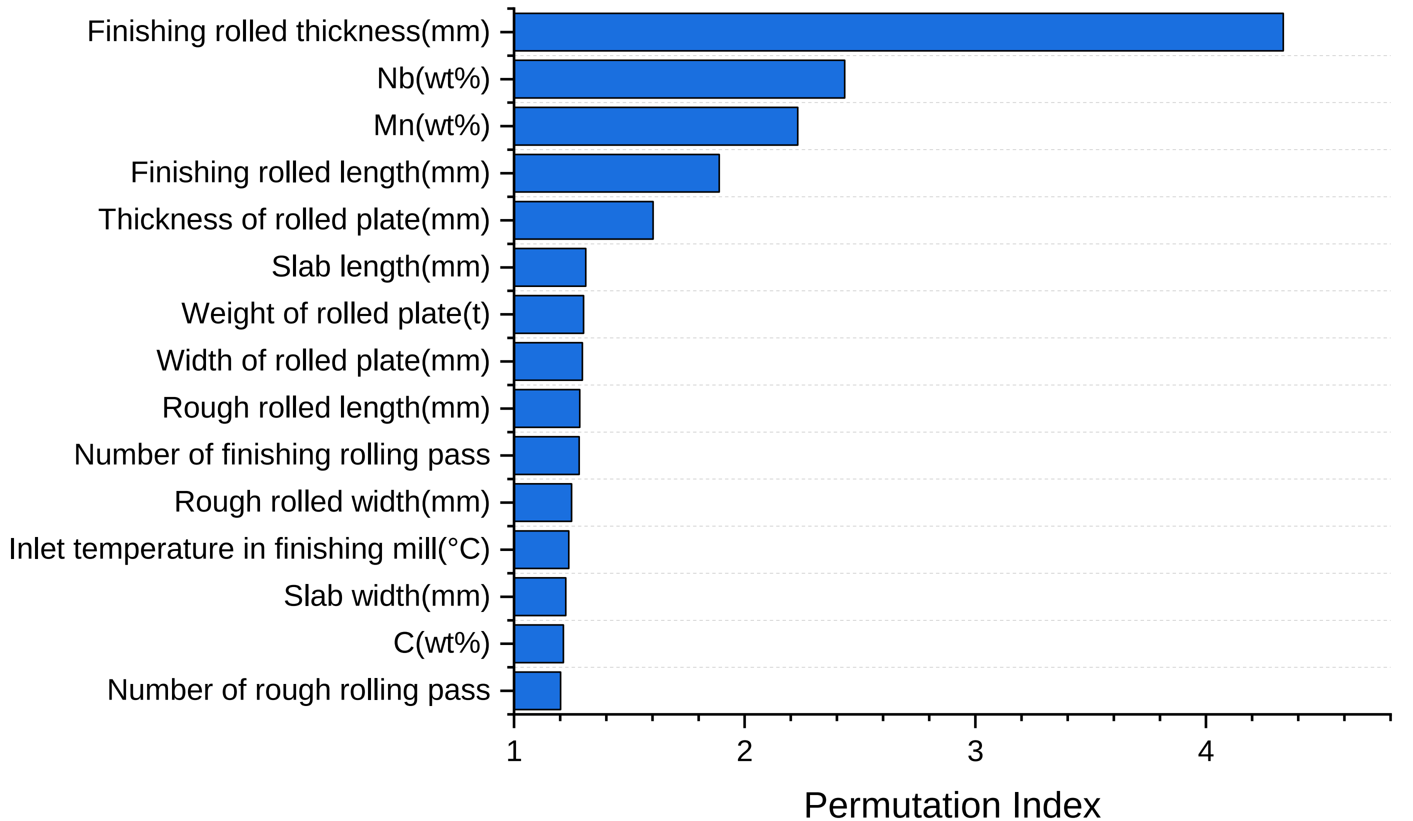


One of the trees



Feature importance ranking

Key influential features ranking based on Neural Networks

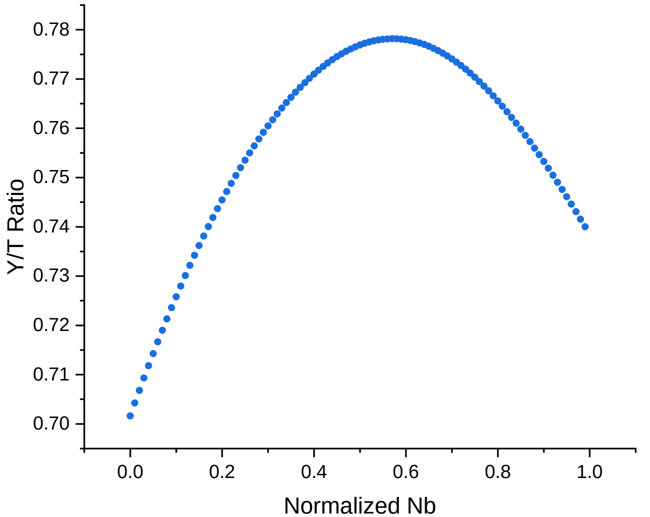


$$\text{Permutation Index} = \frac{MSE_{\ Permutation\ inputs}}{MSE_{\ Original\ inputs}}$$

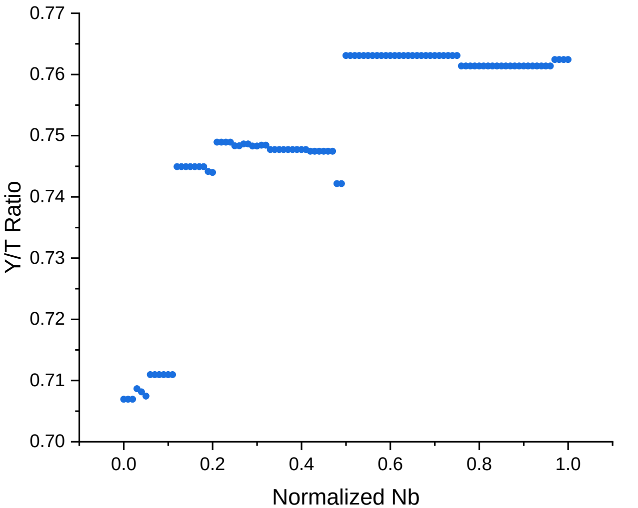
Influence of Mn and Nb

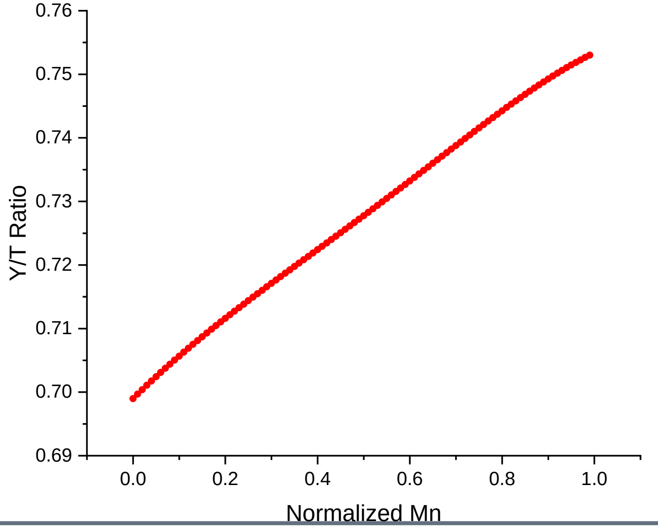
If set input features=mean(except for Mn & Nb):

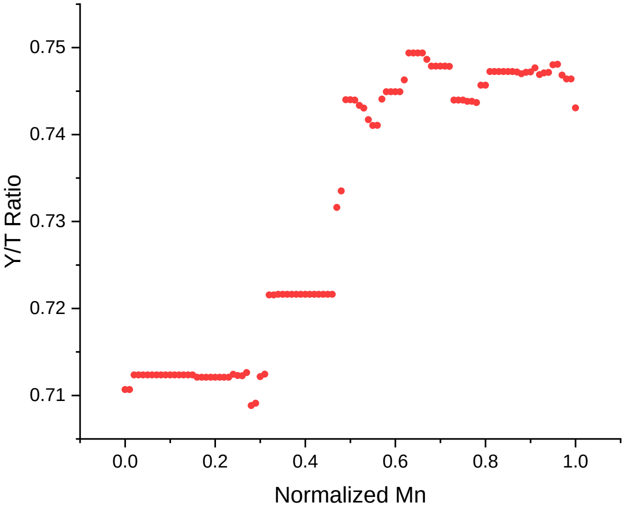
NN :



Xgboost:

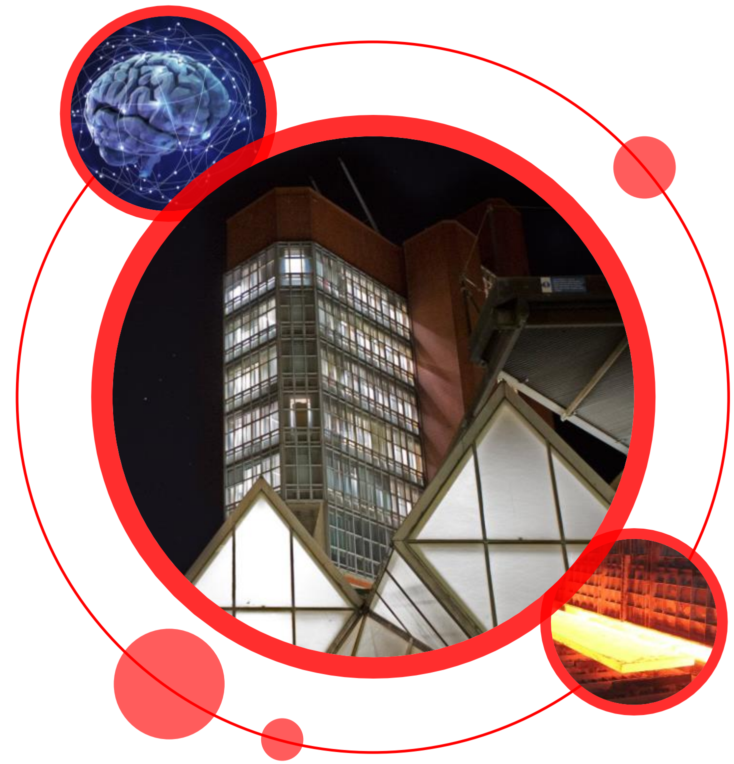






Content

1. Background
2. Prediction of yield to tensile ratio
3. Determination of key influential factors
4. Conclusion



Conclusions & Future work

- We have built models based on machine learning, showing a capability of predicting yield to tensile ratio of hot rolled steel plates with good accuracy

- With the assistance of statistical tools, the model demonstrates the explicit importance of each variable, and recognizes the predominant features on yield to tensile ratio

- Mn and Nb are discovered to be the keys to control the yield to tensile ratio in hot rolling process

Future work

- The specific influence of Mn and Nb on yield to tensile ratio remains unclear and requires further investigation

NISCO

UNIVERSITY OF LEICESTER