Engineering Journal issn: 0125-8281

Article

Predicting the Product Classification of Hot Rolled Steel Sheets Using Machine Learning Algorithms

Chaovarat Junpradub[a] **and Krisada Asawarungsaengkul**[b,*]

Department of Industrial Engineering, Faculty of Engineering, King Mongkut's University of Technology North Bangkok. 1518, Pracharat 1 Road, Wongsawang, Bangsue, Bangkok 10800, Thailand
E-mail: [a]Chaovaratj1973@gmail.com, [b,*]krisada.a@eng.kmutnb.ac.th (Corresponding author)

Abstract. The mechanical properties of the SAPH440 hot rolled steel sheet are mainly controlled to satisfy product specifications. Three mechanical properties including the yield strength, ultimate tensile strength, and elongation are measured and utilized in product classification. Based on these properties, the steel is classified into 3 grades: Class 1 (meets specification), Class 2 (moderate quality), and Class 3 (low). However, various factors can affect the mechanical properties, leading to a long setup time for initial production runs. Therefore, this paper aims to improve the accuracy of these predictions by using machine learning algorithms. The results of experiments showed that the random forest algorithm had the best performance, with an accuracy of 70.0% and a macro average F-1 score of 70.0%. This more accurate prediction can reduce the initial setup time and save 37,000 USD per grade in trial run costs.

Keywords: Hot rolled steel sheet, mechanical properties, product classification, machine learning.

ENGINEERING JOURNAL Volume 27 Issue 8
Received 15 February 2023
Accepted 21 August 2023
Published 31 August 2023
Online at https://engj.org/
DOI:10.4186/ej.2023.27.8.51

DOI:10.4186/ej.2023.27.8.51

1. Introduction

Hot rolled steel is widely used in various industries such as construction, truck manufacturing, shelving, railroads, car parts, machinery, and container production due to its excellent weldability and mechanical properties [1]. Accurate control of mechanical properties such as elongation, yield strength, tensile strength, and impact energy is crucial in the steel industry. As a result, hot strip mills or hot rolling manufacturers focus on predicting and controlling these properties. Traditionally, prediction models have been obtained through simple or multiple linear regression. However, in recent years, machine learning (ML) algorithms have also been used for prediction purposes.

Continuous improvement in hot rolled steel sheet production has been a priority as new products often have specific mechanical property requirements based on their intended application. Currently, process control mainly focuses on the chemical composition of raw materials and rolling conditions. Four critical parameters are adjusted during the process control, including carbon equivalent, thickness, finishing temperature, and coiling temperature. The initial setup involves using scatter plots to determine the relationship between carbon equivalent and mechanical properties, and fine-tuning of process parameters to achieve optimal conditions. This process requires 20 to 25 initial trial runs (each a twenty-ton hot rolled coil) to meet customer specifications, leading to increased production costs and loss of production hours in the hot rolling process.

The hot rolled steel process is illustrated in Fig. 1. The raw material for this process is a slab with a thickness of 220-250 mm. The process consists of the following 8 steps:

1) The slab is reheated in the reheating furnace to a temperature in the range of 1,200 - 1,250 °C,

2) The slab is rolled to reduce the thickness of the transfer bar to about 30 mm. by the roughing mill station,

3) The finishing rolling with transfer bar was passed to the finishing mill at a temperature in the range of 990 - 1,056 °C,

4) Before entering the finishing mill stand, the transfer bar is coiled at the coiled box in order to equalize the temperature of the transfer bar,

5) The transfer bar needed to remove the oxide scale which generates when steel is oxidized at a high temperature,

6) The finishing mill is the process which controls the final thickness of the hot rolled steel sheets by transfer bar changed to steel strip at the finishing mill,

7) The steel strip was cooled to a temperature in the range of 540 - 720 °C by water at the laminar flow before the steel strip is coiled at the down coiler,

8) Hot rolled coil is transferred to the coil yard for cooling to room temperature.

The mechanical properties of the slab after the hot rolling process are altered and must be controlled to meet customer specifications. As new products require tighter specifications for mechanical properties, this work proposes more accurate process control. The main focus of this paper is to predict the quality of the product based on the mechanical properties of a hot strip mill manufacturing process. The linear regression has been utilized to predict the mechanical properties of hot rolled steel sheets, but its results have proven to be inaccurate, leading to high costs during the initial production run. Machine learning algorithms are thus employed to improve the accuracy issue and reduce setup costs.

The following is a review of literature on the influential factors affecting the mechanical properties of hot-rolled steel and the application of prediction algorithms. The use of neural networks to predict the mechanical properties of hot-rolled steel was explored in studies [1, 2]. In particular, study discussed a deep neural network (DNN) model with an R2 value of 0.907 was provide by [1]. The potential for real-time prediction using neural networks was noted as a way to significantly reduce labor and material costs, as detailed in [2].

Matsubara et al. [3] highlighted the importance of mechanical properties to customers. They noted that various factors such as chemical composition, reheating temperature, soaking time of the slab in the reheating furnace, finishing temperature, reduction of thickness, and coiling temperature influence the mechanical properties.

The chemical composition of steel plays a significant role in determining its mechanical properties, with studies indicating that precipitation hardening and grain size strengthening are key factors in solid solution strengthening. The reheating process can also affect the mechanical properties through homogenization of the chemical composition in the slab during soaking. It was further noted that an increase in carbon content can enhance the strength of steel. As such, converting the other alloying elements into carbon terms is a powerful method for predicting the mechanical performance and behavior of steel. The relationship conversion of other alloying elements to carbon is called the carbon equivalent (*Ceq*), as shown in Eq. (1), provided by Agarwal et al [4].

$$Ceq = C + \frac{Mn}{6} + \frac{Si}{24} \tag{1}$$

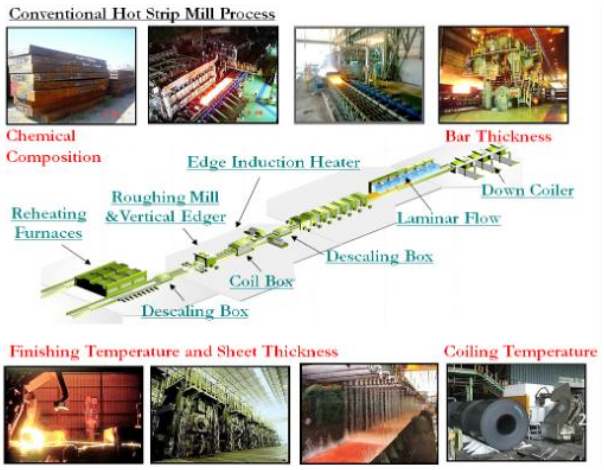

Fig. 1. Hot rolled steel sheet process.

Currently, the regression models are used to predict the mechanical properties of the hot strip rolling process. In quality control, yield strength, tensile strength, and elongation must be accurately predicted and monitored. Steel manufacturers currently rely on the carbon equivalent for prediction; however, it is recognized that hot strip rolling parameters such as rolling temperature can also impact the mechanical properties.

Additionally, the metallurgical structure is very complex. According to study [5], the effect of thermal treatment on microstructure and mechanical properties was investigated. The impact on yield strength, ultimate tensile strength, hardness, and microstructure was examined by considering two factors: soaking temperature and rolling temperature.

It was observed that yield strength, ultimate tensile strength, and hardness could be improved by using the following conditions: a soaking temperature greater than 1100°C, rolling temperatures around 800°C, and a rolling reduction of 50% or more. Rudkins et al. [6] compared finite-element models and the Ekelund and Sims model with actual values to investigate the rolling load of the hot strip mill and found that the predicted values had good agreement with the Ekelund and Sims rolling models.

Junpradub and Asawarungsaengkul [7] discussed papers on predictions using regression models. They found that single linear regression (SLR) and multiple linear regression (MLR) models were used to predict the mechanical properties of hot rolled steel sheets. The SLR model correlated the carbon equivalent with the mechanical properties, while the MLR model added the process parameters to the regression. The results showed that the MLR model was more appropriate than the SLR in predicting yield strength (YS) and elongation (EL), while the SLR model provided more accurate predictions of ultimate tensile strength (TS).

In recent years, machine learning, a component of artificial intelligence (AI), has effectively been applied to various engineering fields. Ghaisari et al. [8] explored its use in the steel industry to predict yield strength (YS), ultimate tensile strength (TS), and elongation (EL) using machine learning. They also used it to identify the crucial input parameters affecting mechanical properties. Ma et al. [9] applied machine learning to diagnose faults in the hot rolling mill process through a semi-supervised coupling fault classification. The study utilized multi-learning (MTL) to classify data from different groups and adaptively predict the mechanical properties of the hot rolling mill process.

Eligüzel et al. [10] studied the use of k-nearest neighbors (k-NN) as predictors on the Twitter platform, finding that k-NN had an accuracy of 0.9 and an F1-score of 0.86 in classification, outperforming support vector machine (SVM) with an accuracy of 0.83 and F1-score of 0.74. Both SVM and k-NN techniques are also explored in Hasan et al. [11]. Ming et al. [12] investigated the hot rod rolling process design using knowledge from the decision tree by random forest (RF) and an ontology design, aimed at optimizing gear box design.

Bagheripoor and Bisadi [13] created a neural network model to accurately predict the rolling force and torque in the hot strip mill process to address nonlinear issues. Lu et al. [14] utilized support vector machine, neural network, and extreme gradient boosting as regression models to predict mill chatter or vibration in cold rolling for ST14 steel. The results showed that extreme gradient boosting achieved the best prediction performance with a high determination coefficient (R^2) and low mean absolute percentage error (MAPE).

This paper focuses on using machine learning to predict product classification. The study includes the following steps: Section 2 outlines the material and methods, including research data, predictions using multiple linear regression (MLR), product classification, data cleaning, methodologies, and model performance measurement. Section 3 determines the optimal hyperparameters of k-nearest neighbors (k-NN), support vector machine (SVM), random forest (RF), and Artificial Neural Network (ANN). In Section 4, the results of model decision-making are validated by using confusion matrix metrics (F1-Score, Precision, Recall, and Accuracy) to evaluate accuracy, effectiveness, and robustness of the classification group of hot rolled coil. The performance of the models is discussed and concluded in Section 5.

2. Material and Method

Currently, the process uses linear regression to predict three mechanical properties (YS, TS, and EL) and product classification. Table 1 lists dome data of input parameters selected as regressors for the regression model. The input parameters consist of carbon equivalent (*Ceq*), slab thickness (THK), finishing temperature (FT), and coiling temperature (CT). However, this approach has limited accuracy in predicting the mechanical properties and product classification.

The linear equations obtained from the regression model are as follows:

$$YS = f_1(Ceq, THK, FT, CT) \quad (2)$$

$$TS = f_2(Ceq, THK, FT, CT) \quad (3)$$

$$EL = f_3(Ceq, THK, FT, CT) \quad (4)$$

These equations have been used as a guideline for process setup and resulted in high wastes in this hot rolled process including: setup time, setup costs, and operation capacity.

2.1. Data for Prediction Model

In this study, the data of hot rolled SAPH440 steel was obtained and analyzed from a laboratory certified by Thai Industrial Standards (TIS 17025-2561) using the test coil methodology according to JIS G3113 for structural automobile hot rolled SAPH440 steel. The data was collected from July 2010 to July 2020. The material was tested using an Instron 5585H series universal testing machine, following the JIS Z2241 standard for ultimate tensile strength of metallic materials. The data used for

training and testing ML models included 7,607 hot rolled sheets. Both mechanical properties (Ceq, THK) and rolling parameters (FT, CT) were considered as factors affecting the controlled mechanical properties. The thermal rolling conditions (FT and CT) are crucial for the yield strength (YS), ultimate tensile strength (TS), and elongation (EL). Table 1 lists some input parameters of the hot rolled sheets used for training the machine learning models.

Table 1. Some of the input parameters of the mechanical properties.

Ceq	*THK* (mm.)	*FT* (℃)	*CT* (℃)
0.29	2.0	841	581
0.28	2.0	841	581
0.26	2.0	864	591
0.26	2.0	864	591
0.30	4.5	858	623
0.30	4.5	850	614
0.27	3.2	875	612
0.29	3.2	875	612
0.28	2.6	875	612

2.2. Product classification

The mechanical properties of SAPH440 hot rolled steel sheet (YS, TS, and EL) is used to determine the product classification of hot rolled steel sheet. Scores are assigned based on these properties: "Fail" (YS < 315, TS < 460, EL < 31), "Medium" (315 < YS < 365, 460 < TS < 510, 31 < EL < 35), and "Good" (YS > 365, TS > 510, EL > 35). Table 2 summarizes the criteria for each score. The quality of SAPH440 hot rolled steel sheet is classified into three categories based on the scores of YS, TS, and EL (as shown in Table 3). The criteria for each category are:

• Class 1 (Preferable): All YS, TS, and EL scores meet the "Good" requirement.

• Class 2 (Moderate): At least one YS, TS, or EL score is "Medium".

• Class 3 (Low): At least one YS, TS, or EL score is "Fail".

Table 2. The scores for the mechanical properties of the products.

Mechanical properties	Good	Medium	Fail
Yield Strength	315 - 365	> 365	< 315
Ultimate Tensile Strength	460 - 510	> 510	< 460
Elongation	> 35	31 - 35	< 31

Table 3. Product Classification.

Yield Strength	Tensile Strength	Elongation	Class
Good	Good	Good	1 (Preferable)
Medium ≥ 1 for 3 properties			2 (Moderate)
Fail ≥ 1 for 3 properties			3 (Low)

2.3. Data Cleaning

To train and test the ML model, data was collected from 7,607 hot rolled sheets over the past 10 years. The unbalanced data set from the hot rolling process is shown in Fig. 2. Outliers were removed during the analysis. To balance the data, equal sample sizes of hot rolled steel sheets were randomly selected from each of the three product quality classes. The resulting balanced data set consisted of 1,500 data points per class. This balanced data set ensured accuracy and precision in prediction.

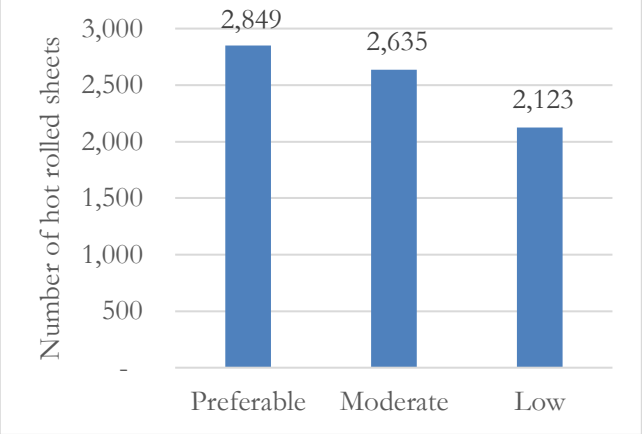

Fig. 2. Data set for test and train process of ML.

2.4. Methodologies

This paper uses machine learning algorithms to improve decision-making in the hot rolling process and achieve a preferable class of product quality. Machine learning can be divided into three sub-fields: supervised learning, unsupervised learning, and reinforcement learning. This study focuses on supervised learning for classification and prediction. Supervised learning involves finding the relationships between input features and the output data set to improve prediction accuracy.

The k-nearest neighbors (k-NN), support vector machine (SVM), random forest (RF), and artificial neural network (ANN) algorithms were tested to find the best performing one. The machine learning models were trained using input features (CEQ, THK, FT, CT) and the actual class of SAPH440 hot rolled steel sheets. The framework for product classification prediction is shown in Fig. 3.

Data set was typically divided into three groups including training, test, and validation. The recommended

value for the proportionality between the three data groups was described in [15-20]. In this paper, the data set of 4,500 samples were divided into 70% for training and 30% for testing during creating machine learning models.

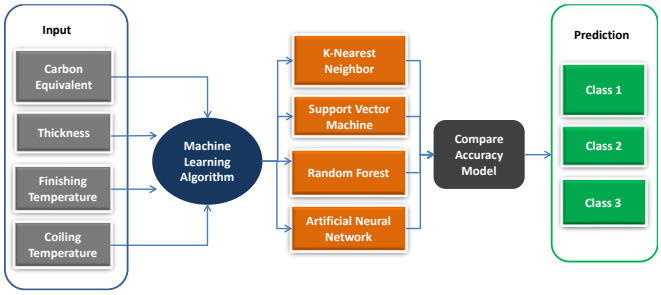

Fig. 3. Methodology for classification and prediction.

2.4.1. k- Nearest Neighbors (k-NN)

The k-Nearest Neighbors (k-NN) algorithm is used for supervised learning classification and regression. In the classification method, it compares input data distances to classify a data point into a group. The k-NN algorithm chooses the k nearest data points, where k is any integer. It then uses a voting system for the most frequent class in case of classification or averages the labels in case of regression. The performance and classification method of k-NN are described in [21-22]. The Euclidean distance between two points can be calculated using Eq. (5). The k-NN algorithm predicts a new data point's class by using the number of nearest neighbors, as shown in Fig. 4, and classifying it based on the Euclidean distance.

$$\text{Euclidean} = \sqrt{\sum_{i=1}^{k}(X_i - Y_i)^2} \qquad (5)$$

where Euclidean is the distance between new data point and *i* point; k is number of neighbors.

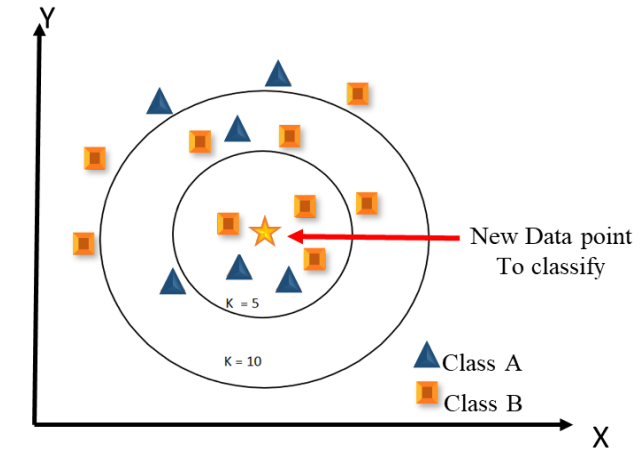

Fig. 4. The architecture of k-nearest neighbors (k-NN).

2.4.2. Support Vector Machine (SVM)

The support vector machine (SVM) is a non-linear learning algorithm as illustrated in Fig. 5. The objective of SVM is to project the dataset into a feature space and construct a hyperplane that separates the data into different classes. The data points laid closest to the hyperplane are called support vectors, and the maximum distance between the hyperplane and these data points is the margin. The hyperplane will result in higher accuracy of classification if it can produce a maximum margin. A larger margin means lower generalization error. Support vector machine can work well in high-dimensional data when the number of input features is greater than the number of observations or samples.

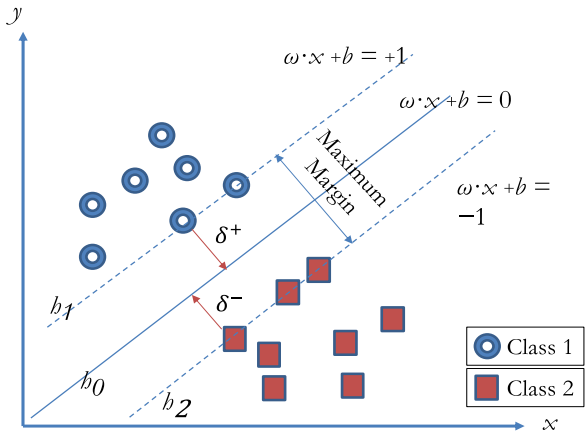

Fig. 5. The architecture of the support vector machine (SVM).

The prediction model design and techniques to minimize the least square error between the real data and prediction model output are discussed in [23]. SVM also has a kernel function that can transform complex data into a linearly separable form in a higher dimension. The classification function for SVM is explained in Eq. (6) below:

$$y = f(x) = \omega^T x + b \qquad (6)$$

where x is set of input sample features, y is the output result, ω is set of weights for each feature, b is a bias.

The Support Vector Machine (SVM) uses an optimization approach to maximize the margin between data points. This helps reduce the number of weights that are non-zero and contributes to the important features. The margin is the sum of the shortest distances to the closest positive and negative points, denoted as δ^+ and δ^-, respectively. This leads to the optimal hyperplane being found. The non-zero weights are linked to the support vectors, which are the data points closest to the hyperplane.

Hyperplanes as in Fig. 5 can be defined by Eq. (7) and Eq. (8).

$$\omega \cdot x + b = +1 \quad when, y = +1 \qquad (7)$$

$$\omega \cdot x + b = -1, when\ y = -1 \qquad (8)$$

The median hyperplane is displayed in Eq. (9).

$$\omega \cdot x + b = 0 \tag{9}$$

2.4.3. Random Forest (RF)

Random forest is a supervised machine learning algorithm. Random forest (RF) is a powerful machine learning model for classification and regression problems with a strong predictive ability. It works by combining multiple decision trees to create a "forest". In classification models, each decision tree acts as a classifier, and the final prediction is made based on the majority vote of the decision trees, as shown in Fig. 6. There are three types of nodes in a random forest: root nodes, internal nodes, and leaf nodes. Root nodes are the first branching point in a decision tree, internal nodes represent additional decisions based on the root node or other internal nodes, and leaf nodes are the final outputs of the decision tree.

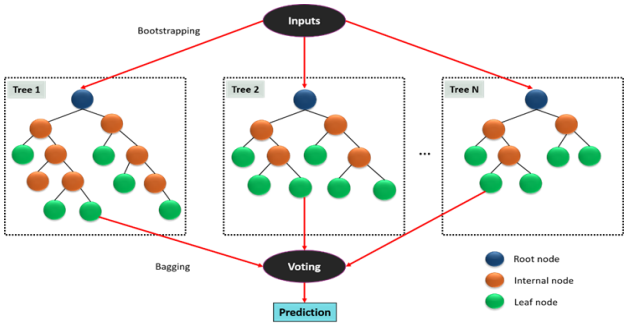

Fig. 6. The architecture of the Random Forest for classification.

In the tree diagram of random forest, each module represents the significance of an attribute value among instances, leading to branches and possible input values to meet the criteria. To classify an instance, the process begins with creating a root node, determining the attribute data, and branching the tree based on the desired attribute value. The accuracy of the prediction model is influenced by the design of the input parameters, including both primary and secondary input features, as shown in [24].

Since the algorithm process method was considered repeated, the additional sub-tree root node was defined as the new node. RF uses the "bagging" method in which the sampling method is utilized to combine the learning models that can improve the prediction result and reduce the over-fitting. This formula employs a decision tree which is used to determine the concept of how nodes produce an effect on a decision tree branch, and can be expressed in Eq. (10). This concept is called the Gini index.

The RF itself in another way is labelled by the entropy in order to encourage nodes to branch in a decision tree, as depicted in Eq. (11). The decision entropy was applied by statistical probability in order to select how the node should branch, which is measured by using the Gini index. Entropy provides the measurement of impurity in a group of observations. Entropy of a dataset closed to zero means that it is not useful for learning.

$$\text{Gini} = 1 - \sum_{i=1}^{c} (P_i)^2 \tag{10}$$

$$\text{Entropy} = \sum_{i=1}^{c} - P_i * log_2 (P_i) \tag{11}$$

where P_i is probability of class i
c is the number of classes

2.4.4. Artificial Neural Network (ANN)

The artificial neural network (ANN) is modeled after the human brain and has complex learning processes. It typically consists of three layers: input, hidden, and output. The input layer receives data, which is then processed by the hidden layer through an activation function and sent to subsequent layers as seen in Fig 7. There are various activation functions, such as Softmax, Leaky ReLU, Sigmoid, Tanh, and ReLU. The ANN used in this paper is a feed-forward network, which processes data from the input layer through the hidden layer and outputs the result through the output layer. The artificial neural network can be used for various problems, as referenced in [15-16, 25-26].

The backpropagation algorithm is a key component in fine-tuning the weights of an artificial neural network (ANN). It calculates the gradient of the loss function for a single weight using the chain rule. Optimal weights reduce error rates and improve the accuracy of the ANN's predictions. The use of ANN in predicting roll force and roll torque in the hot strip rolling process is demonstrated in [25]. The relationship between inputs and outputs is calculated through a weighted sum of input features and bias, as shown in Eq. (12). The output result is generated through the ANN's feed-forward process, starting from the input layer and flowing through the hidden layer to the output layer is display in Fig 8.

$$y = \delta(\sum_i \omega_i x_i + b) \tag{12}$$

where, y is the output
ω_i is the weight coefficient i
x_i is input feature i
b is the bias vectors

Equation (13) is used for calculating a output of a hidden layer that use information from the previous layer and then send the value to the next layer.

$$y^l = \delta(\omega^l y^{l-1} + b^l) \tag{13}$$

where, y^l is the output of layer l
ω^l is the weight coefficient matrix of layer l
b^l is the bias vectors of layer l
$\delta(z)$ is the activation functions.

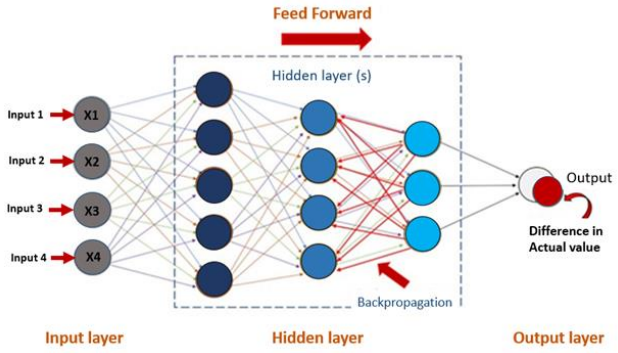

Fig. 7. The feedforward Neural Network Algorithms.

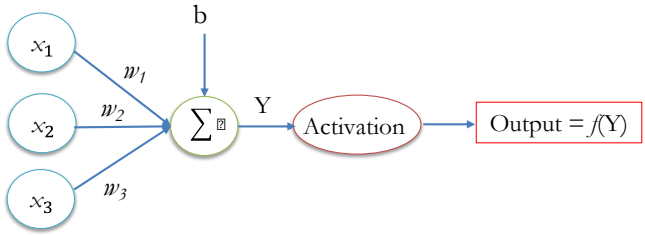

Fig. 8. The process of computing the output result for ANN.

The data was normally divided into three groups: training and testing, with recommended proportions described in [15-20]. In this paper, a dataset of 4,500 samples was split into 70% for training and 30% for testing during fitting the machine learning model. The ANN architecture design involves three layers: input, hidden, and output. Optimal hyperparameters can improve the accuracy of machine learning models.

2.4.5. Performances of prediction models

Efficiency of classification is an important methodology in terms of evaluation based on the measured accuracy, precision, recall and F1-Score. The formulars to measure the performance of model are shown in Eq. (14) to (17). The parameters represent the number of True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (14)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (15)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (16)$$

$$\text{F1-score} = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (17)$$

The confusion matrix is a tool used in classification problems to visualize the performance of a classifier. The proposed method of correction using a confusion matrix evaluates the accuracy of the classifier by comparing the actual classes to the predicted classes, with positive and negative classes defined. Fig. 9 provides an illustration of how correct and incorrect predictions can be counted when the target class is class 1.

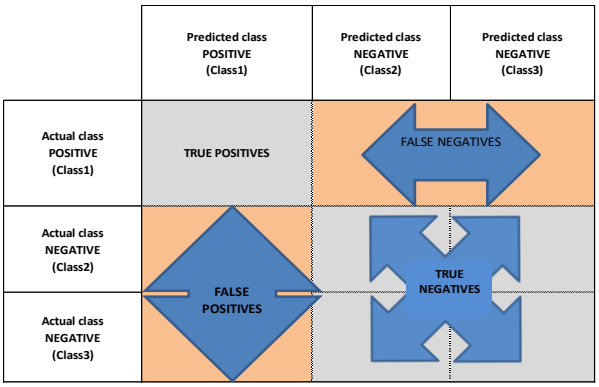

Fig. 9. The confusion matrix platform when Class 1 is the target class.

F1-Score is the weighted average of precision and recall, two important metrics in classification problems. Precision (also known as positive predictive value) measures the accuracy of positive classifications, calculated as the ratio of True Positives to Total Predicted Positives (Eq. (15)). Precision is appropriate when the cost of False Positives is high. Recall (also known as sensitivity) measures the proportion of actual positive cases correctly identified, calculated as the ratio of True Positives to Total Actual Positives (Eq. (16)). Recall is relevant when the cost of False Negatives is high. F1-Score balances precision and recall (Eq. (17)) and is suitable for classification problems with unbalanced class distribution.

3. Hyperparameters for ML algorithms

The experiments were done to determine the suitable hyperparameters for ML algorithms used for product classification. The optimal hyperparameters can enhance the prediction accuracy of ML models.

3.1. k-Nearest Neighbors (k-NN)

The k-NN algorithm classifies a new sample by taking a majority vote of its k nearest neighbors, determined by a distance function. The algorithm is trained using input features and actual class outputs, which in this paper are Class 1, 2, and 3. The goal of the experiment is to predict the product classification in the hot rolled process for new sample data. The optimization of the hyperparameters in the k-NN learning process involves adjusting the value of k, as shown in Table 4.

Table 4. Hyperparameters for k-NN.

Parameter Keys	Range
n_neighbors	5
algorithm	auto
leaf_size	30
metric	minkowski
metric_params	None
p	2
weights	uniform

The best n_neighbors value for maximum accuracy is 5. The k-NN model's accuracy in classifying hot rolled products is 67%.

3.2. Support Vector Machine (SVM)

In this paper, the support vector machine (SVM) is utilized for classification problems. The hyperplanes are built to make the product classification. The optimal hyperparameters based on the training dataset for SVM are shown in Table 5.

Table 5. Optimal hyperparameters for SVM.

Parameter Keys	Range
C	1000
gamma	0.0001
cache_size	200
max_iter	1
Probability	FALSE
Shrinking	TRUE
decision_function_shape	ovr
degree	3
kernel	rbf
verbose	FALSE

SVM model with Radial Basis Function (RBF) kernel is successfully improved the efficiency with an accuracy of 57.0% for classification. The RBF kernel (K) can be expressed as in Eq. (18):

$$K(x_1,x_2) = \exp\left(\frac{\|x_1 - x_2\|^2}{2\sigma^2}\right) \quad (18)$$

where σ is the variance and our hyperparameter; $\|x_1 - x_2\|$ is the Euclidean distance between two points $x_1$and x_2.

3.3. Random Forest (RF)

Table 6 displays the optimal hyperparameters for the Random Forest model. The n_estimators, which is the number of trees created before voting, is found to be optimal at 100. The recommended value for max_depth, the maximum number of levels in each decision tree, is 3. Additionally, the optimal max_feature, the maximum number of features considered in each tree, is "auto".

Table 6. Optimal hyperparameters for random forest.

Parameter Keys	Range
n_estimators	100
max_feature	auto
max_depth	3
Random state	42
criterion	entropy

The initial dataset was divided into 70% training samples and 30% testing samples for use by the Random Forest model in making classifications. This resulted in a classification accuracy of 70%.

3.4. Artificial Neural Network (ANN)

The product classification of the hot rolled process was performed using a single-layer Neural Network (ANN) architecture. The model consists of three layers: the input layer with four features (chemical content and hot rolled parameters), the hidden layer with 4 nodes and 16 hidden layers, and the output layer with 3 neurons for the three classes of hot rolled steel sheet. The optimal hyperparameters for the ANN, including the number of hidden layers and nodes, as well as the recommended hyperbolic tangent activation function, are shown in Table 7.

A common logistic model for activation function used for multi-class classification is sigmoid, $S(x)$. The sigmoid activation function can be expressed as in Eq. (19)

$$S(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{1+e^x} \quad (19)$$

where, $S(x)$ = sigmoid function, e = Euler's constant, equal to 2.7182.

According to the definition of the mathematical function, the S-curve is a Sigmoid function also known as the logistic function, which is crucial in providing the optimal minimum error for linear functions. The Sigmoid function maps real numbers to a range between 0 and 1, and its purpose is to convert real values from inputs into probabilities for the outputs.

For the classification of the hot rolled steel sheet, the hyperbolic tangent (tanh) activation function, shown in Eq. (20), is recommended. The relationship between the sigmoid and hyperbolic tangent functions is expressed in Eq. (21). The tanh activation function has an s-shaped curve and ranges from -1 to 1.

$$F(x) = \tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})} \quad (20)$$

$$\tanh(x) = 2 \cdot S(2x) - 1 \quad (21)$$

The Stochastic Gradient Descent (SGD) algorithm is used to minimize the objective of the loss function. After training, the recommended solver for the ANN is the Adam optimizer, an extension of the SGD technique, for the supervised learning model. A learning rate of 0.1 is also suggested.

Table 7. Optimal hyperparameters for ANN.

Parameter Keys	Range
hidden_layer_size	(16,4)
activation	Tanh
solver	Adam
learning_rate_int	0.1
max_iter	(1000, step = 50)

The accuracy of the product classification performed using the Artificial Neural Network (ANN) was found to be only 40%, which is considered to be low. This result raises the need for further investigation and improvement to increase the accuracy of the model.

4. Results and Discussions

Four machine learning algorithms were employed to develop a prediction model for the product classification of hot rolled steel. The hyperparameters for these algorithms were determined in a previous section. The performance of the experiment on these ML models is reported in this section, using precision, recall, accuracy, and F1-score as the evaluation metrics. These metrics are displayed in Fig. 10 to 13. Accuracy is a commonly used performance indicator for selecting the best ML algorithm, while the F1-Score, a combination of precision and recall, is an appropriate metric to use when dealing with unbalanced class distributions.

The results from the validation experiments reveal that the accuracy of the k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN) algorithms are 67%, 57%, 70%, and 40%, respectively. It can be observed that the Random Forest (RF) algorithm provides the best accuracy for all classes of classification. Furthermore, in terms of precision, recall, and F1-score, the RF algorithm outperforms the k-NN, SVM, and ANN algorithms. The performance of the ANN was found to be the worst among all the algorithms in terms of all performance metrics. As a result, the Random Forest algorithm is selected as the best option for predicting in the hot roll process in manufacturing due to its superior performance compared to the other algorithms.

In addition to the machine learning algorithms, the existing regression model that is currently used in the manufacturing process was also employed to predict the product classification. Based on the dataset used for modelling these machine learning algorithms, the accuracy of the regression model was only 33.5%. The regression model tends to predict the product quality as Class 2. It is found that 4,487 out of 4,500 hot rolled sheets were predicted as Class 2. True positive (TP) for Class 2 is 1,500 out of 1,500 hot rolled sheets. That means number of False Positives (FP) for Class 2 is 2,987 out of 3,000 hot rolled sheets. The reason that why performance of regression model is very low because it tries to predict the mean of mechanical properties which are "Medium".

It can be concluded that the machine learning algorithms are more accurate in comparison to the regression model. Among the machine learning algorithms, the Random Forest algorithm demonstrated the best performance and thus, is considered to be the most suitable for the product classification task.

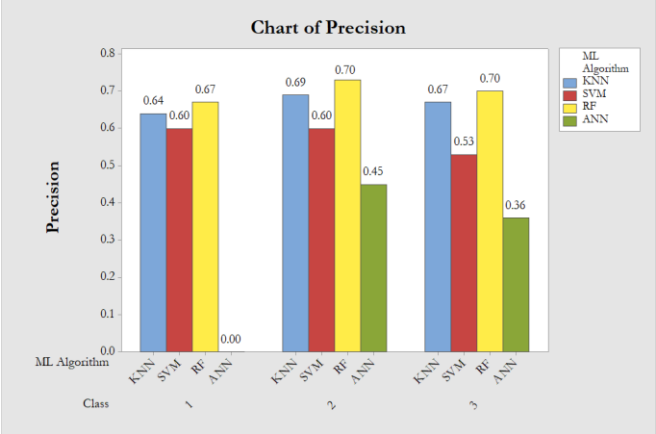

Fig. 10. The precision value of ML algorithms.

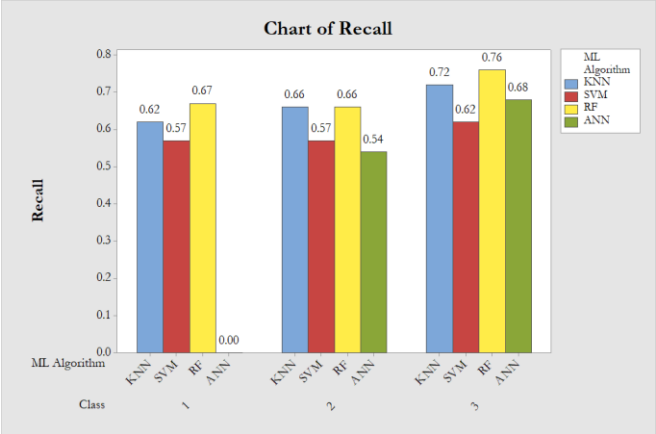

Fig. 11. The recall value of ML algorithms.

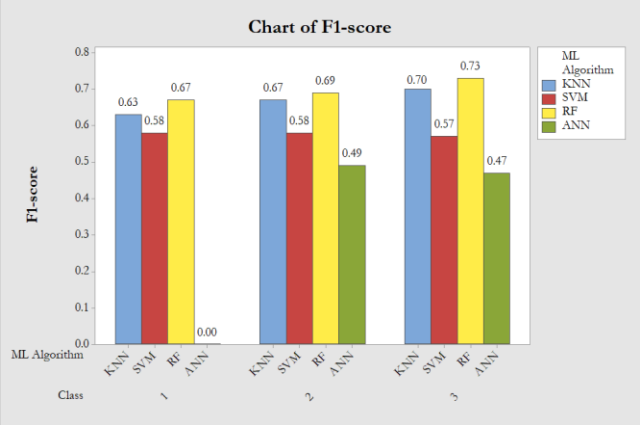

Fig. 12. The F1-score value of ML algorithms.

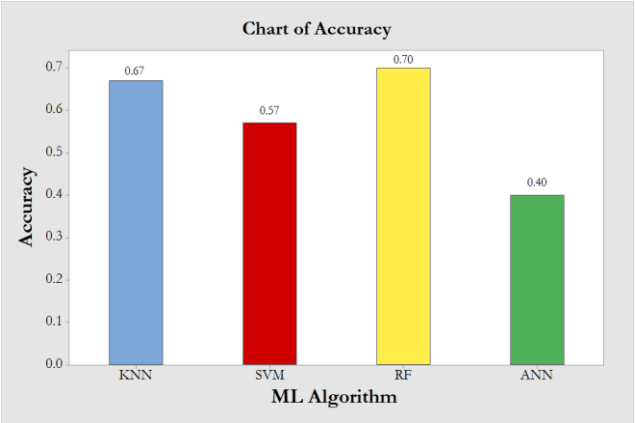

Fig. 13. Compare accuracy of ML algorithms.

5. Conclusion

The product quality classification in the manufacturing process is based on three physical properties: tensile strength, ultimate tensile strength, and elongation. Previously, the linear regression method was used to predict the product quality. However, this method showed low accuracy, leading to longer initial setup times and higher setup costs. Additionally, misclassification of product quality led to increased financial losses for the manufacturing process.

To address this issue, new prediction models using machine learning algorithms were proposed. The results of the experiments showed that these algorithms could significantly improve the accuracy of product quality prediction compared to the linear regression model. Among the algorithms tested, random forest was found to have the best performance, scoring highest in all performance metrics.

The use of the selected machine learning algorithm model not only improves the accuracy of product quality prediction but also reduces setup time and costs. It is estimated that the benefits of the prediction model could lead to a reduction of approximately 37,000 USD in trial products, development time, and trial costs per grade.

The further research opportunity is to develop more accurate prediction model by utilizing the importance input features and the hybrid machine learning model. The closen-loop control system and optimization technique can be developed to control the hot rolled process so that this process will be able to produce more preferable product quality.

References

[1] Q. Xie, M. Suvarna, J. Li, X. Zhu, J. Cai, and X. Wang, "Online prediction of mechanical properties of hot rolled steel plate using machine learning," *Mater. Des.*, vol. 197, p. 109201, 2021.

[2] V. V. Kurban, N. L. Yatsenko, and V. I. Belyakova, "Feasibility of using neural networks for real-time prediction of the mechanical properties of finished rolled products," *Metallurgist*, vol. 51, no. 1–2, pp. 3–6, 2007.

[3] H. Matsubara, T. Osuka, I. Kozasu, and K. Tsukada, "Optimization of metallurgical factors for production of high strength, high toughness steel plate by controlled rolling," *Trans Iron Steel Inst Jap*, vol. 12, no. 6, pp. 435–443, 1972.

[4] F. Kawamura, M. Miura, R. Ebara, and K. Yanase, "Material strength of long-term used penstock of a hydroelectric power plant," *Case Studies in Structural Engineering*, vol. 6, pp. 103-114, 2016.

[5] S. N. Prasad and D. S. Sarma, "Influence of thermomechanical treatment on microstructure and mechanical properties of a microalloyed (Nb + V) weather-resistant steel," *Mater. Sci. Eng. A*, vol. 399, no. 1–2, pp. 161–172, 2005.

[6] N. Rudkins and P. Evans, "Mathematical modelling of mill set-up in hot strip rolling of high strength steels," *J. Mater. Process. Technol.*, vol. 80–81, pp. 320–324, 1998.

[7] C. Junpradub and K. Asawarungsaengkul, "Mathematical modeling to predict the mechanical properties of hot rolled steel sheets," in *Proc. Int. Conf. Ind. Eng. Oper. Manag.*, 2019, pp. 2636–2644.

[8] J. Ghaisari, H. Jannesari, and M. Vatani, "Artificial neural network predictors for mechanical properties of cold rolling products," *Adv. Eng. Softw.*, vol. 45, no. 1, pp. 91–99, 2012.

[9] L. Ma, K. Peng, J. Dong, and C. Hu, "A novel semisupervised classification framework for coupling faults in hot rolling mill process," *ISA Trans.*, vol. 111, pp. 376–386, 2021.

[10] N. Eligüzel, C. Çetinkaya, and T. Dereli, "Comparison of different machine learning techniques on location extraction by utilizing geo-tagged tweets: A case study," *Adv. Eng. Informatics*, vol. 46, p. 101151, 2020.

[11] M. S. Hasan, A. Kordijazi, P. K. Rohatgi, and M. Nosonovsky, "Triboinformatic modeling of dry friction and wear of aluminum base alloys using machine learning algorithms," *Tribol. Int.*, vol. 161, p. 107065, 2021.

[12] Z. Ming, G. Sharma, J. K. Allen, and F. Mistree, "An *Ontology for Representing Knowledge of Decision Interactions in Decision-Based Design*," *Comput. Ind.*, vol. 114, p. 103145, 2020.

[13] M. Bagheripoor and H. Bisadi, "Application of artificial neural networks for the prediction of roll force and roll torque in hot strip rolling process," *Appl. Math. Model.*, vol. 37, no. 7, pp. 4593–4607, 2013.

[14] X. Lu et al., "Prediction and analysis of cold rolling mill vibration based on a data-driven method," *Appl. Soft Comput. J.*, vol. 96, p. 106706, 2020.

[15] T. S. Wang, C. H. Tan, L. Chen, and Y. C. Tsai, "Applying artificial neural networks and remote sensing to estimate chlorophyll-a concentration in water body," in *Proc. - 2008 2nd Int. Symp. Intell. Inf. Technol. Appl. IITA 2008*, vol. 1, pp. 540–544.

[16] Y. S. T. Hong, "Dynamic nonlinear state-space model with a neural network via improved sequential learning algorithm for an online real-time hydrological modeling," *J. Hydrol.*, vol. 468–469, pp. 11–21, 2012.
[17] A. Bayram, M. Kankal, and H. Önsoy, "Estimation of suspended sediment concentration from turbidity measurements using artificial neural networks," *Environ. Monit. Assess.*, vol. 184, no. 7, pp. 4355–4365, 2012.
[18] A. Bayram, M. Kankal, G. Tayfur, and H. Önsoy, "Prediction of suspended sediment concentration from water quality variables," *Neural Comput. Appl.*, vol. 24, no. 5, pp. 1079–1087, 2014.
[19] H. Tabari and P. H. Talaee, "Reconstruction of river water quality missing data using artificial neural networks," *Water Qual. Res. J. Canada*, vol. 50, no. 4, pp. 326–335, 2015.
[20] M. Zounemat-Kermani, Ö. Kişi, J. Adamowski, and A. Ramezani-Charmahineh, "Evaluation of data driven models for river suspended sediment concentration modeling," *J. Hydrol.*, vol. 535, pp. 457–472, 2016.
[21] R. F. Ivo et al., "Study and classification of the crystallographic orientation distribution function of a non-grain oriented electrical steel using computer vision system," *J. Mater. Res. Technol.*, vol. 8, no. 1, pp. 1070–1083, 2019.
[22] L. B. Marinho, P. P. Rebouças Filho, and V. H. C. de Albuquerque, "Ultrasonic sensor signals and self organized mapping with nearest neighbors for the microstructural characterization of thermally-aged Inconel 625 alloy," *Comput. Ind.*, vol. 107, pp. 1–10, 2019.
[23] S. Bouhouche, L. Laksir Yazid, S. Hocine, and J. Bast, "Evaluation using online support-vector-machines and fuzzy reasoning. Application to condition monitoring of speeds rolling process," *Control Eng. Pract.*, vol. 18, no. 9, pp. 1060–1068, 2010.
[24] Y. Xu, B. Zheng, and M. Zhang, "Capacity prediction of cold-formed stainless steel tubular columns using machine learning methods," *J. Constr. Steel Res.*, vol. 182, p. 106682, 2021.
[25] J. Ghaisari, H. Jannesari, and M. Vatani, "Artificial neural network predictors for mechanical properties of cold rolling products," *Adv. Eng. Softw.*, vol. 45, no. 1, pp.91–99, 2012.
[26] W. Cheevachaipimol, B. Teinwan, and P. Chutima, "Flight Delay Prediction Using a Hybrid Deep Learning Method," *Eng. J.*, vol. 25, no. 8, pp. 99-112, Aug. 2021.

Chaovarat Junpradub is a Ph.D. candidate of Department of Industrial Engineering, Faculty of Engineering, King Mongkut's University of Technology North Bangkok. He got the Bachelor Degree in the Field of Production Engineering from Faculty of Engineering, King Mongkut's University of Technology North Bangkok, Thailand. He earned the master degree of Industrial engineering from Chulalongkorn University, Thailand. His experience in field of quality assurance and laboratory test. He was one of the team that developed the hot rolled steel for automotive industry in Thailand. His expertise in the quality assurance and product development in hot rolled steel sheets. For now, he tries to develop the precise prediction model for hot rolled steel sheets for automotive industry.

Krisada Asawarungsaengkul is an Associate Professor and head of Operations Research and Engineering Management Research Center in Department of Industrial Engineering, Faculty of Engineering, King Mongkut's University of Technology North Bangkok. He holds the B.Eng. in Production Engineering from King Mongkut's University of Technology North Bangkok, M.Eng. in Industrial Engineering from Chulalongkorn University, and the Ph.D. in Industrial Engineering from Sirindhorn International Institute of Technology, Thammasat University, Thailand. Dr. Asawarungsaengkul has also been certified with lean six sigma master black belt. His research interests include lean six sigma, manufacturing improvement, optimization, operations research, supply chain management, and lean office and service.