

# Processing Technique Selection for Steels Based on Mechanical Properties Using Machine Learning Framework

Amitava Choudhury (✉ [a.choudhury2013@gmail.com](mailto:a.choudhury2013@gmail.com))

University of Petroleum and Energy Studies <https://orcid.org/0000-0001-5153-4449>

---

## Research Article

**Keywords:** Steels, Machine learning, Random Forest, Rolling Process, Material Informatics

**Posted Date:** April 19th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-336843/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Processing Technique Selection for Steels Based on Mechanical Properties Using Machine Learning Framework

Amitava Choudhury<sup>1</sup>

<sup>1</sup>School of Computer Science, University of Petroleum and Energy Studies, Dehradun India.

Email: a.choudhury2013@gmail.com

## Abstract

Prediction of process route from materials with a desired set of property is one of the fundamental issues in the perspective of the materials design aspect. The parameter space is often too large to be bound since there are too many possibilities. However, in the areas with limited theoretical access, artificial learning techniques can be attempted by using the available data of the candidate materials. In this study, a computational method has been proposed to predict the different process routes of steel constituted taking composition and desired mechanical properties as inputs. First of all, historical data of the actual rolling process was collected, cleaned, and integrated. Further, the dataset is divided into four different classes based on the rolling process data. Then, to find out the essential characteristics of variables, feature correlation among various features has been calculated. A state-of-art machine learning prediction methods such as logistic regression, K- nearest neighbor, Support vector machine, and the random forest are studied to implement the prediction model. In order to avoid the overfitting of the model, k fold cross-validation is applied to the model, and achieve a realistic prediction result with an accuracy of 97%. The F1-score of the classification model is 0.86, and the kappa score is 0.95, which comply that the model has excellent learning and speculation ability and the precise forecast of steel process routes based on the given input parameters.

**Keywords:** Steels; Machine learning; Random Forest; Rolling Process, Material Informatics

## 1. Introduction

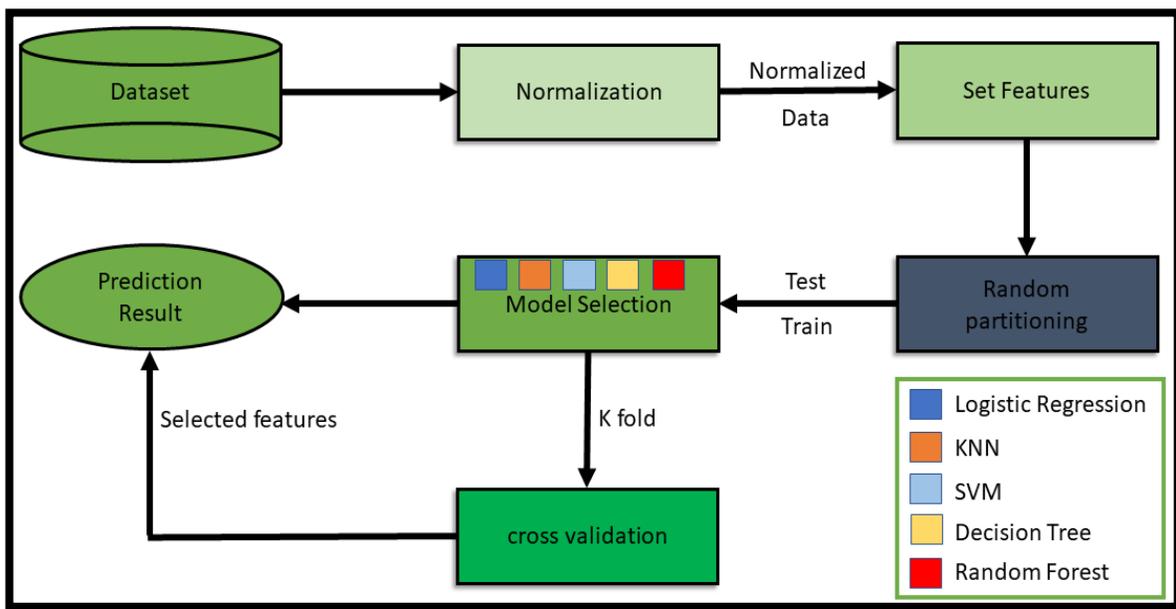
Steel is a versatile material concerning its application due to its superior properties, available variations, and amicability of its shaping and processing to obtain desired application-specific products [1-2]. In case of different grades of steel, many properties depend on composition and microstructure. They can be divided into two major categories, namely structure sensitive properties (e.g., yield strength) and structure insensitive properties (e.g., electrical conductivity) [3]. The material processing depends on various conditions like rolling techniques (hot rolled and cold drawn), which affect the overall performance and application of steel. The cold-rolled process increases the

hardness and decreases the grain size, whereas the hot rolled process reduces the average grain size of metal but maintains an equiaxed microstructure [4-5]. Control of process variables of thermo-mechanical processing can be useful to realize the desired grain size through some solid-solid reactions like recrystallization and phase transformation [6-13]. To reduce the production cost and the human intervention, the introduction of automation in the steel industry in a more enhanced form than the present one, is necessary. In this perspective, prediction algorithms are used to speed up the production process, and such efforts are being taken up by the materials science and technology community for attaining certain solutions after compilation of extensive datasets of material composition, process methods, and related properties [14-16]. In 2009, Brahme *et al.* have designed an artificial neural-network-based prediction model of cold rolling textures from steel, which is used to predict fiber texture using texture intensities, carbon content, carbide content, and the amount of rolling reduction [17]. Simecek *et al.* have developed the MECHP tool to predict the mechanical properties of hot rolled steel products. The machine measures process data like water-cooling and subsequent air-cooling of the hot-rolled narrow plate and wire [18]. Jassim *et al.* have developed an empirical model to predict hardness, yield strength, and tensile strength for single roll melt spinning of rapidly solidified ribbons [19]. This forecasting technique assists with examining the connection between rapid solidification parameters (opening breadth, a gap in the Nozzel-roll wheel and melting temperature) and the thickness, hardness, and elasticity of quick rapid solidified ribbons [19]. Stiffness, strength, ductility, hardness, toughness, *etc.* are some of the mechanical properties, which govern the selection of a material for a given application. Looking for the optimum fusion of material properties along with economic factors, is a necessity for selecting process routes. The determination of the material processing method is of considerable significance while determining the performance of steel. This is due to the fact that variation in process routes for a given grade of steel is capable of producing different structures. This, in turn, is capable of influencing properties like fracture and corrosion behavior [20-21]. It is, therefore, imperative to formulate a distinction between the fundamental and desirable input parameters (here composition, strength, and rolling process). During the manufacturing process, materials and process parameters are to be controlled, and hence, these are ideally desired to be input variables for the study of determining the process routes [22]. The extent of control that the material experiences during processing defines the final properties of the material. At present, a large number of research initiatives have been taken up to predict the steel processing mechanism along with predicting the mechanical properties of steel using various

computational methodologies based on artificial neural network and pattern recognition [23-24]. However, such a complex network structure requires more time and training data to ensure that the model predicts with a high degree of accuracy. In recent trends, many industries are inclined towards machine learning orientated solution to speed up the processing mechanism [25-31]. In this study, a robust machine learning model is described for prediction of different manufacturing processes of steel of various grades viz. hot rolled (HR), cold drawn (CR), annealed cold drawn (ACR), spheroidized annealed cold drawn (SCR). The proposed computational model will be applied without utilizing materials or energy for production. This will make the process less time consuming and thereby economically more viable. This eventually will cater to the needs of the clients of steel industry by lowering the overall production cost.

## 2. Computational Model

In this paper, the effect of various classification models over steel dataset has been discussed to automate the steel processing mechanism [33]. To implement the prediction model, open source software Python 3.7 has been used. The libraries used in this model are Scikit-learn, pandas, and seaborn for various computational and visualization purposes [32]. Figure 1 shows the predictive steel processing model step.



**Figure1.** Schematic representation of steel processing predictive model.

The data used in the trained and testing process includes 11 columns indicating different features and 129 rows indicating the type of steel as mentioned in the supplementary document. Various compositions of steel like Iron (Fe), Carbon (C), Manganese (Mn), Sulphur (S), Phosphorus (P), and various mechanical properties like tensile strength( $\sigma_{TS}$ ), reduction in area ( $\delta_A$ ), hardness ( $H$ ), elongation ( $\delta_E$ ) and yield strength( $\sigma_{YS}$ ), are used as features. We also take one more function named carbon equivalent (CE), which can be evaluated using equation 1.

$$CE = C + \frac{Mn}{6} + \frac{(Cr + Mo + V)}{5} + \frac{(Ni + Cu)}{15} \quad (1)$$

Here  $C = \sum C$  in feature set

$Mn = \sum Mn$  in feature set

$Cr=0, Mo=0, V=0, Ni=0, \text{ and } Cu=0$

In machine learning, logistic regression is useful while obtaining a particular class or event existing from the actual class table. Various classification techniques such as logistic regression, support vector machine, K-nearest neighbor, and random forest is used to compare the results [34-38]. In this paper, a multi class classification is discussed to predict the thermomechanical processing routes, namely hot rolled (HR), cold drawn (CR) annealed cold drawn (ACR), and spheroidite annealed cold drawn (SCR) as describe in table 1.

**Table1.** Various steel processing methods used for the prediction model.

Sl. No.	Steel processing method	Quantity
1	Hot rolled	61
2	Cold Drawn	47
3	Annealed, Cold Drawn	9
4	Spheroidite Cold Drawn	12

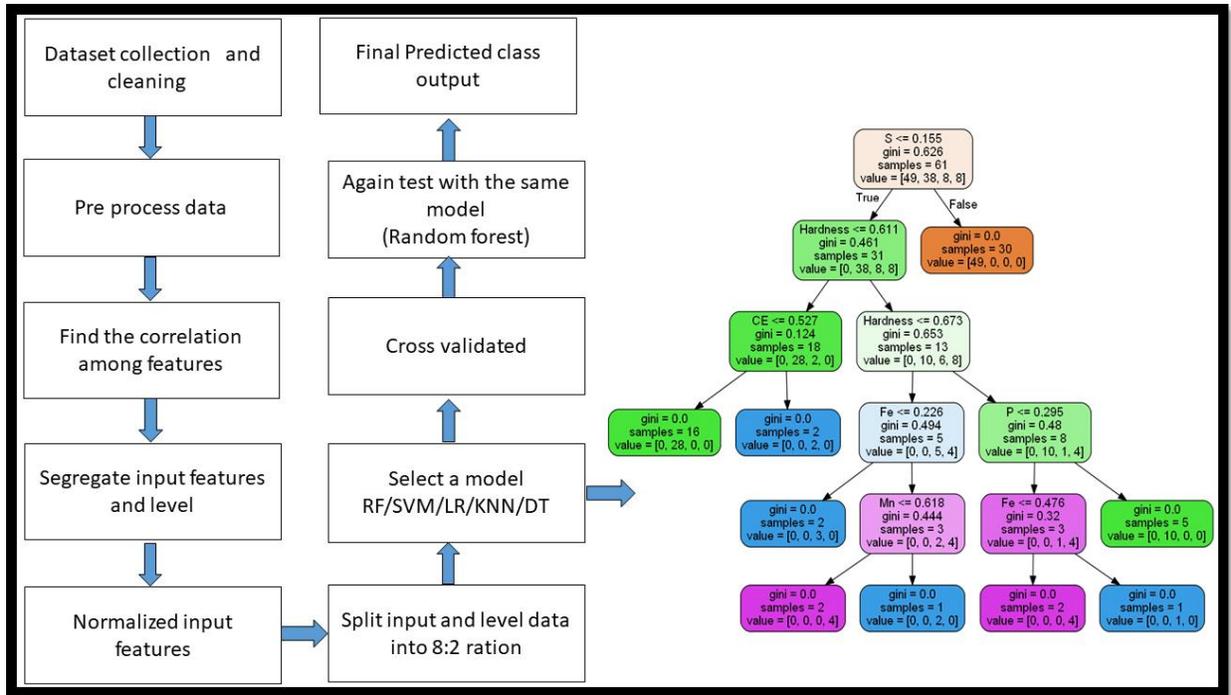
Table 2 displays the statistics summary of each column from the dataset used, which includes, mean, standard deviation, minimum and maximum values of the column in the dataset. There are columns having a count of 129, which shows that the dataset does not have any NULL values. This description also shows that the features of the dataset have varying ranges. Machine learning models like logistic

regression perform better when they learn from data that is in the same range and to accomplish that normalization is a necessity for this dataset.

**Table2:** Descriptive statistics of input variables

	<b>Fe</b>	<b>Me</b>	<b>C</b>	<b>CE</b>	<b>S</b>	<b>P</b>	$\sigma_{TS}$	$\sigma_{YS}$	$\delta_E$	$\delta_A$	<b>H</b>
<b>count</b>	129	129	129	129	129	129	129	129	129	129	129
<b>mean</b>	0.48	0.53	0.37	0.47	0.29	0.17	0.49	0.50	0.25	0.45	0.48
<b>std</b>	0.21	0.24	0.23	0.23	0.32	0.25	0.23	0.26	0.28	0.22	0.23
<b>min</b>	0	0	0	0	0	0	0	0	0	0	0
<b>max</b>	1	1	1	1	1	1	1	1	1	1	1

According to the statistical distribution of attributes in Table 2, each attribute requires some tweaking. The collected data is in different range and if the scales for different features are wildly different, it may cause the abnormality in model performance. To ensuring all features weighted equally in their representation, data normalization should be used to adjust their weights accordingly. Data normalization is a procedure frequently applied as a major aspect of data preparation for AI. The objective of normalization is to change the estimations of numeric sections in the dataset to a typical scale, without twisting contrasts in the scopes of qualities. This seems like a better choice for more realistic data representation. Figure 2 shows the overall flow of the proposed work. Besides, it shows how a tree can be generated from the used dataset. We tested the data with various machine-learning algorithms, and we found bagging based random forest gives the best results after parameter tuning.



**Figure 2.** Process steps involved for the model development. The tree in the right side shows how data slit into the left and right node so that optimal results can be derived by using a voting scheme.

### 2.1.Feature Correlations

In predictive statistics and machine learning, an attribute with a high correlation coefficient has more influence on the prediction variable [36]. This can be understood and visualized with the help of a correlation map. The correlation coefficient ( $r$ ) can measure the relationship between variables. The correlation coefficient of two attributes is always a range between 1 (Positive relationship) to -1 (Negative relationship), whereas 0 implies no correlation at all. Equation 2 is useful to calculate the correlation coefficient:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (2)$$

The correlation map displayed in Figure 3 shows the dependence that the features have on each other. The correlation coefficients are presented within the cells. The diagonals possess a correlation coefficient of 1, showing that each feature perfectly correlates with itself. Yield strength has strong negative correlations with elongation and reduction in the area. In contrast, elongation has strong

positive correlations with a reduction in area ( $\delta_A$ ), and strong negative correlations with yield strength ( $\sigma_{YS}$ ), hardness ( $H$ ), and tensile strength ( $\sigma_{TS}$ ).

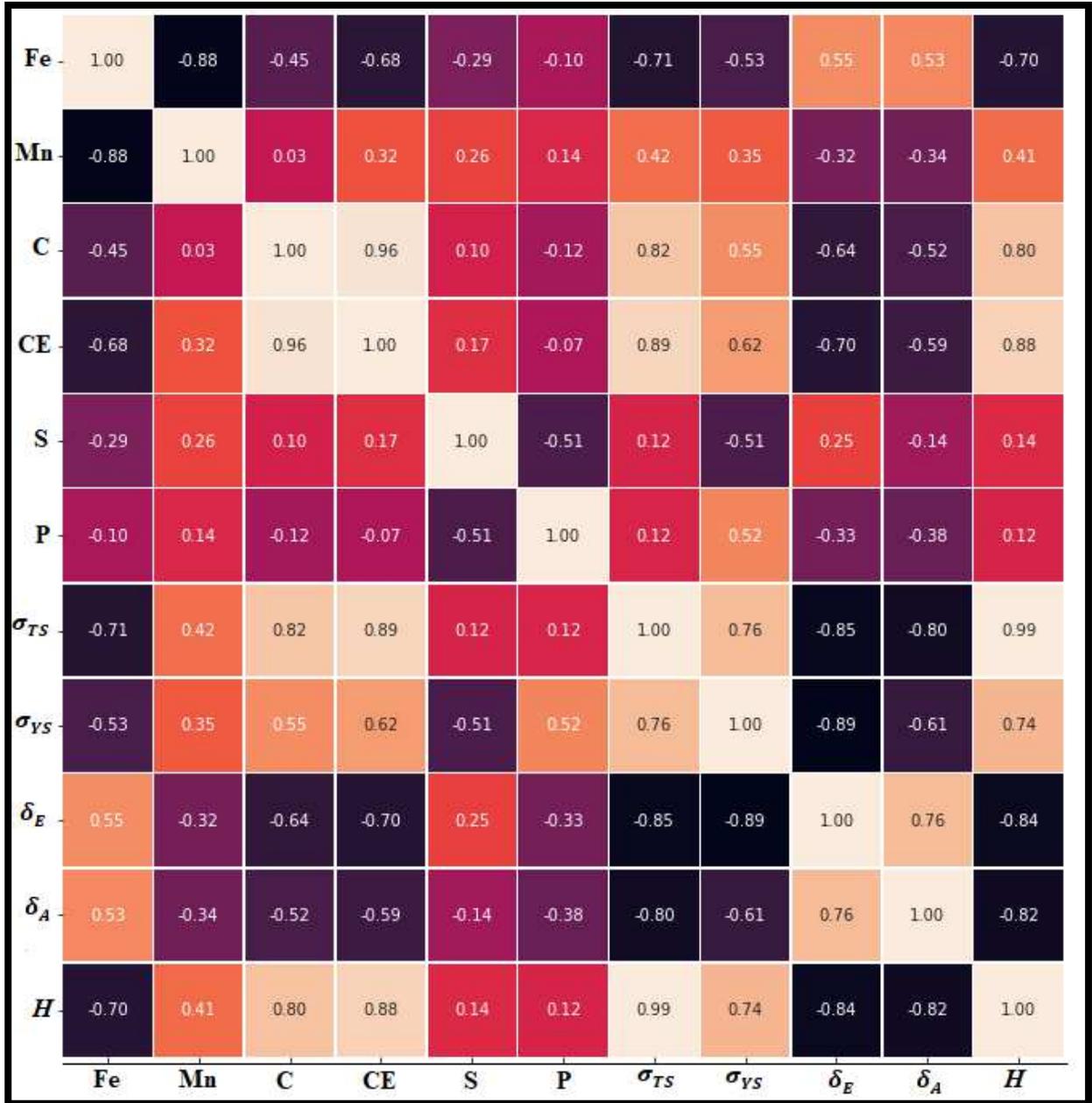
## 2.2. Random Forest implementation

Classifier such as ‘Decision Tree’ often encounters limitation on the complexity at a fundamental level [39]. When it comes to unsupervised learning, decision trees’ can grow to arbitrary complexity by achieving generalized accuracy only. On the other hand, the oblique decision tree is beneficial for having optimized training set accuracy. Hence, the proposed method is to increase the number of trees in feature subspace using random criteria, forming a random forest. By increasing the number of trees, the complementary generalization in the combined classification can be monotonically improved.

Using the best split approach, the above-selected features are used for finding the root node of the tree. With the establishment of a root, the child nodes computed and determined using the same approach as earlier. Finally, the process will halt only when the root node and target as a leaf node are set. In the same way, all the possible trees with the ‘k’ feature selection to form ‘n’ random trees of the forest. When it comes to validation of any feature set against the expected outcome, a very stable test set should be built. In a classification method such as random forest, a large set of combinations is generated, which will increase the complexity in the outcome as well. Hence, a voting methodology is adopted that can determine the maximum occurrences of the target output from the available ‘n’ random trees. So, of the different criteria available for the selection of the target node, majority voting is adopted. This method tries to cast votes of all the possible target feature selections available from the random forest. The feature set that has maximum votes obtaining for maximum test cases is considered as winner feature set for the given classification, as shown in the following equation.

$$f = \frac{1}{N} \sum_{k=1}^N f_n(x') \quad (3)$$

Where,  $N$  = total number of features,  $k$  = number of features selected,  $f_n$  = classification tress generated,  $x'$  = input complements provided.



**Figure 3.** Correlation map: X and Y-axis parameters are Iron, Manganese, Carbon, Carbon equivalent, Sulfur, Phosphorous, Tensile Strength, Yield strength, Elongation, Reduction in area, Hardness.

### 2.3.Support Vector Machine

In machine learning, the support vector machine (SVM) is a statistical enabled learning method based on the strategy of the highest margin. This method is appropriate when the sample set is lesser than the number of dimensions. The main working principle of SVM is creating hyperplane to uniquely

classified data points from an n-dimensional space where n is the number of features. An SVM is a soft margin classifier, and the meaning of this is that we will have hyperplane and margin around the hyperplane, but not all the observations need to be on the correct side of the hyperplane. Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane can be attributed to different classes. [40].

In the linearly separating case, the decision surface equation of the separating hyperplane can be written [41] as

$$W^b + x = 0 \tag{4}$$

Where, x is input vector, w is an adjustable weight vector and b is bias.

#### **2.4. Decision tree**

Decision trees are a popular supervised learning method that, like many other learning methods, can be used for both regression and classification [42-43]. The working principle of decision trees is to split the data into subsets, where each subgroup belongs to only one class. This is accomplished by dividing the input space into a pure plane (region). Practically in real life, extremely pure subsets are not possible; therefore, to get real data points, the data is too divided into many subsets in such a manner where each subset belongs from the same class.

Boundaries uniquely identified in each section are called decision boundaries based on which the decision tree model makes classification decisions. A decision tree is a hierarchical structure with nodes and directed edges. The top node is known as the root node, and the bottom is known as leaf nodes. Nodes that are neither the root node nor the leaf nodes are called internal nodes. The root and internal nodes have test conditions; each leaf node has a class label associated with it. Figure 5 represents a clear idea about different nodes. A classification decision is made by traversing the decision tree starting with the root node.

The depth of a decision tree is the number of edges in the longest path from the root node to the leaf node. The decision algorithm can operate on the original training data pretty much as is. So decision trees tend to work well with data sets that have a mixture of feature types binary, categorical, or continuous and with features that are on very different scales. One drawback of decision trees is that despite the use of pruning, they can still overfit all or parts of the data and may not achieve the best generalization performance compared to other methods.

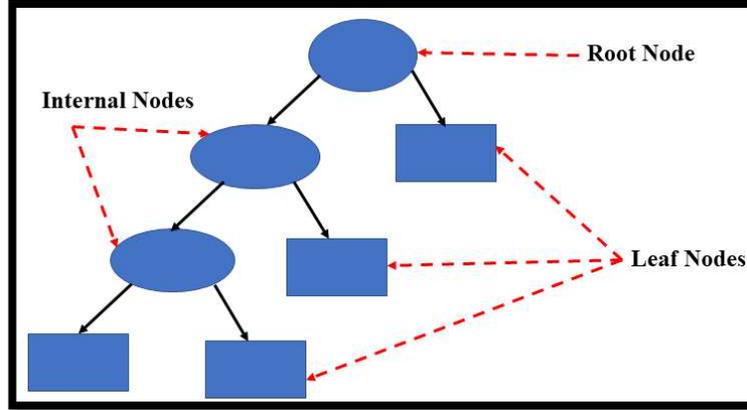


Figure 4. Node block diagram of a decision tree. Left-hand side nodes are known as the left child node, and the right side nodes are known as the right child node.

With the training dataset  $\{x_i, Y_i\}$ , the approximation function can be expressed as [42]:

$$F_k(x) = F_{k-1}(x) - \gamma_k h_k \quad (5)$$

Where, the corresponding training dataset of decision tree  $h_k$  is  $\left\{x_i, -\frac{\delta L}{\delta \hat{y}_i} [F_{k-1}(x_i), y_i]\right\}$  and  $y_k$  is  $\arg \min \sum_{i=1}^n L(y_i, F_{k-1}(x_i) + \gamma_k h_k)$ .

### 2.5.K Nearest Neighbor

kNN is one of state of the art supervised machine learning classification algorithm. The idea behind kNN is to classify data points from its neighbor. KNN relies on the notion of the so-called duck test. In the classification problem, data points from the same class are identified. Hence, similar input samples labeled with the same target label. This means that classification of a sample is dependent on the target labels of the neighboring points.

kNN value can be measured by calculating the distance between its neighbors, and such distance can be calculated using Euclidean, Manhattan, Minkowski equation. If  $K = 1$ , at that point, the case is mostly assigned to the class of its closest neighbor. This distance calculated using, as given below equation

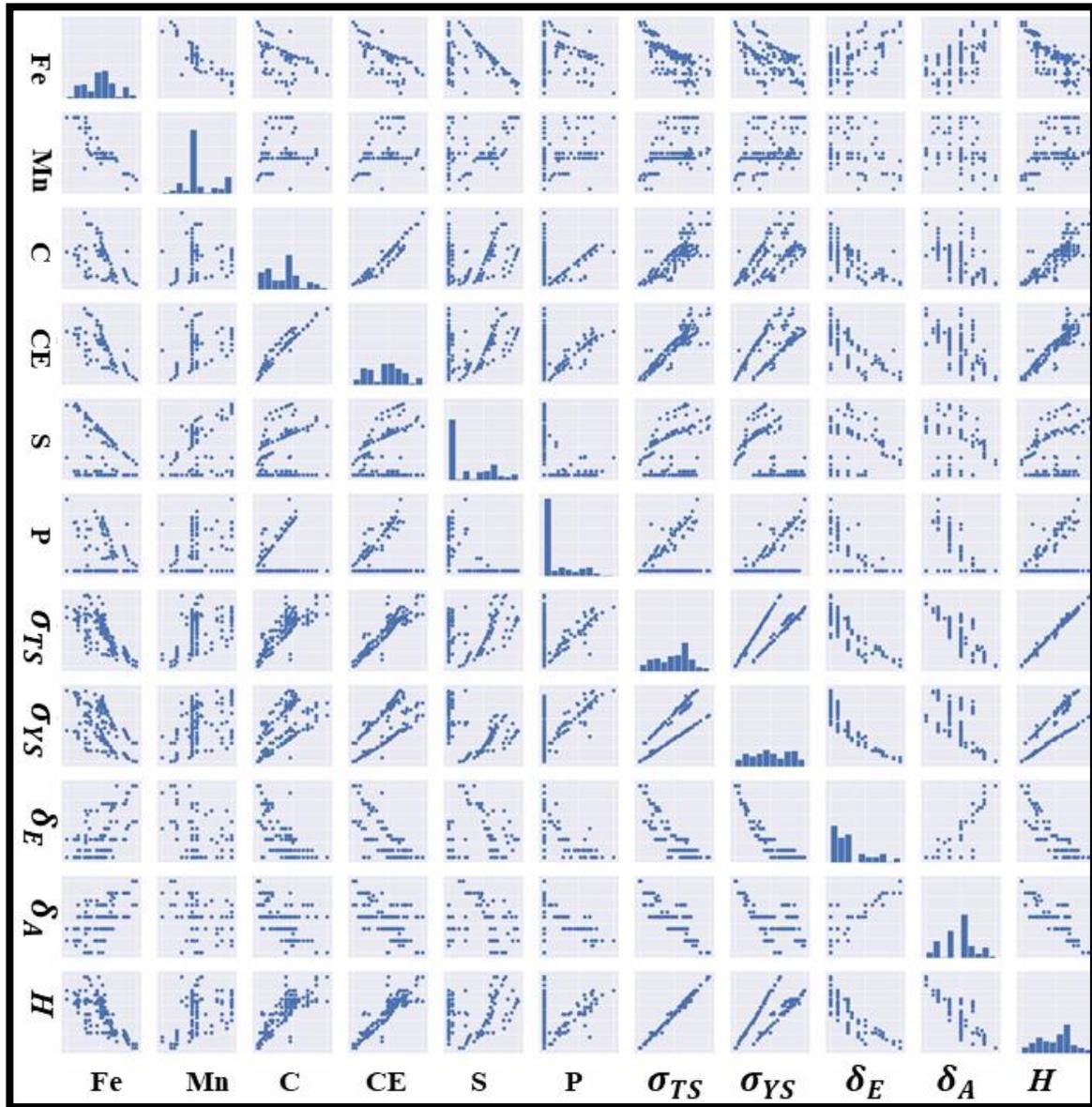
Euclidian:  $\sqrt{\sum_{i=1}^k (x_i + y_i)^2}$

Manhattan:  $\sum_{i=1}^k |x_i - y_i|$

Minkowski:  $(\sum_{i=1}^k (|x_i - y_i|^q))^{1/q}$

### 3. Results and Discussion

Finding the correlation among various compositions and mechanical properties is central to comprehend the connection between those features with other features..



**Figure 5.** Correlation among various features using a scatter plot. The diagonal matrix indicates the highest correlation, among other features.

The input dataset is on different scales; hence it is not good practice to use such a dataset without using data preprocessing; therefore, we normalize all data to maintain a specific range of input data samples by using min-max feature scaling.

$$\acute{x} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (6)$$

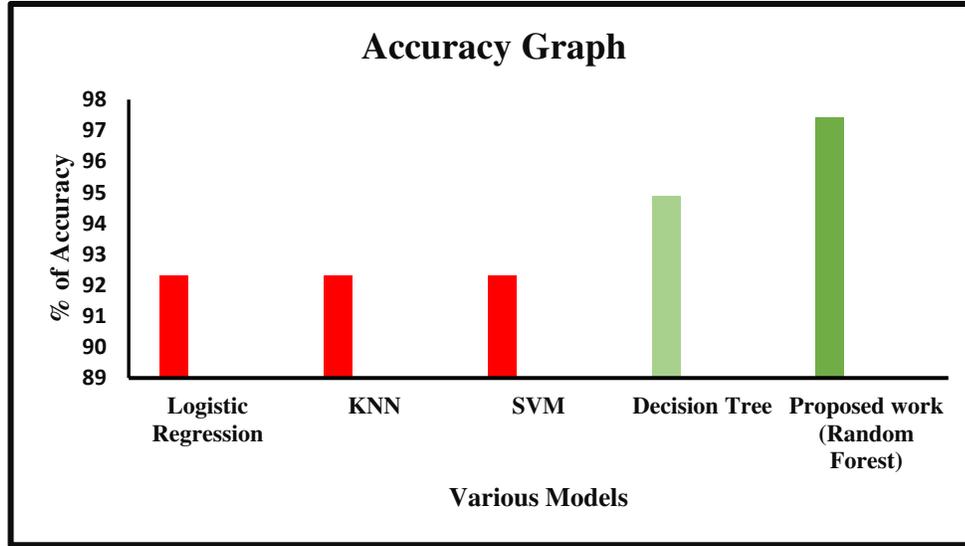
Where  $x$  is original and  $\acute{x}$  is normalized value.

While computing the performance of the model, the relationship among various compositions and mechanical properties is, need to understand—correlation among compositions and mechanical properties graphically represented in figure 5 using the python seaborn module. From the scatter plot, it is shown that when yield strength increase the value of elongation decreases and while increasing the elongation, hardness of the surface decreased, which is evident to the ground truth theory. This correlation scatters plot provides a quantitative description of all input parameters, which is very useful in this study. It is also identified that not all input parameters are played a vital role in the correlation process, so it can be removed from the feature set to reduce the number of input parameters in such a manner that will not affect the results. In this paper, a multi-class classification is discussed; namely, hot rolled, cold drawn, annealed cold drawn and spheroidite annealed cold drawn are denoted as class 1, class 2, class 3 and class 4 respectively.

The dataset split into two subsets, 80% of data is used for training, and 20% is used for testing. In this study, a computational comparison has been discussed among various existing machine-learning algorithms such as logistic regression, Decision tree, k nearest neighbor, SVM, and random forest, as shown in table 3. A graphical representation of the accuracy table is shown in figure 6.

**Table3:** Comparison table of the model prediction accuracy with various classification techniques.

	<b>Accuracy Obtained (%)</b>
Logistic Regression	92.30
KNN	92.30
SVM	92.30
Decision Tree	94.87
<b>Proposed work (Random Forest)</b>	<b>97.43</b>



**Figure 6.** Comparisons of the accuracy of different machine learning models.

### 3.1.Feature (predictive properties) selection

Classification accuracy is not enough to measure the effectiveness of any model. Based on precision and recall, the f1 score can be calculated, which is a statistical measure of testing accuracy. The F1 score value near 1 indicates the perfect precision and recall. F1 score can be calculated as equation 7

$$F1 - Score = 2 * \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

Where,  $Precision = \frac{TP}{TP+FP}$  and  $Recall = \frac{TP}{TP+FN}$

$TP=$  True positive;  $FP=$  False Positive,  $FN=$  False Negative.

On the other hand, a kappa score (k) is also used to measure the test accuracy quality. The significance of rater reliability quality lies in the way that it represents the degree to which the information gathered in the testing process are right portrayals of the factors estimated.

Kappa score 0 means the the model understanding isn't appropriately mapped and 1 represent flawless understanding, and this value can be calculated as

$$k = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e} \quad (8)$$

Where,  $p_o$  : relative observed agreement among raters

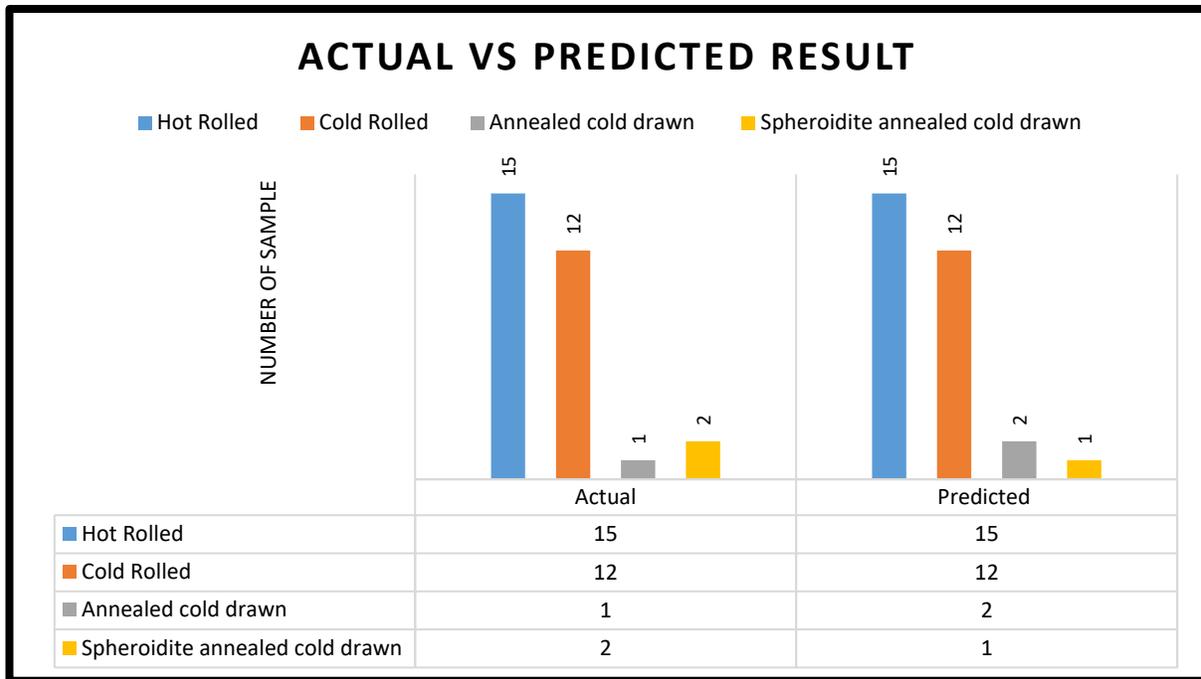
$p_e$  is the hypothetical probability of chance agreement.

Table 4 shows the model accuracy, precision, recall, and kappa score on the used dataset.

**Table 4:** Confusion Matrix of the proposed model. The matrix shows the number of samples for each class predicted by the model.

Predicted Class labels	Actual Class labels				Class
	hot rolled	cold drawn	annealed cold drawn	Spheroidite annealed cold drawn	Precision
hot rolled	20	0	0	0	100
cold drawn	0	15	0	0	100
annealed cold drawn	0	0	1	0	50
Spheroidite annealed cold drawn	0	0	1	2	100
<b>Class Recall</b>	100	100	100	66.66	<b>Total Accuracy</b>
					<b>97.43</b>
				<b>F1-Score</b>	<b>0.8666</b>
				<b>Kappa score</b>	<b>0.95</b>

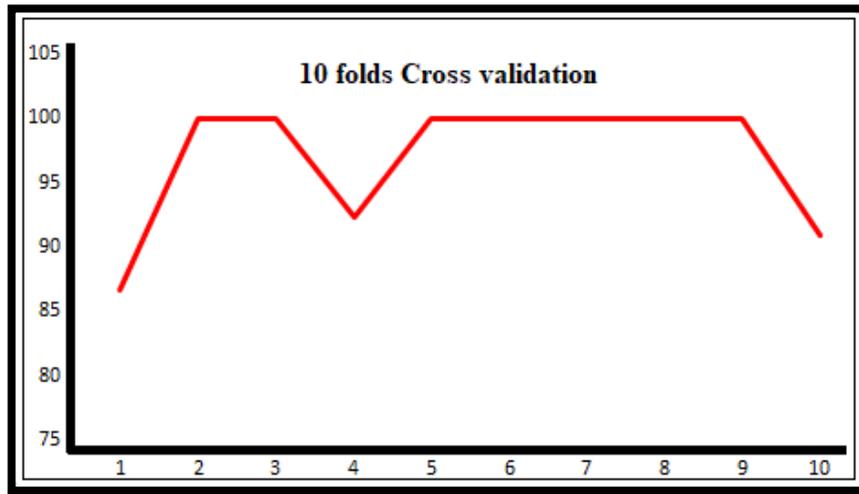
*F1* score is used to understand the correctness of classification if the score is one that indicates the model is stable and 0 means unstable. In this study, we achieve ~0.97% accuracy, which shows the goodness of the model. Moreover figure 7, represents a clear picture of the performance of testing data. From figure 6, it can be inferred that except in one case, all predicted results are correctly classified; hence, such a model can be useful in automating the process selection process in the steel industry. Number of input parameters are also played a vital role in model performance and a model could be more effective if fewer input parameters are used to achieve the desire results. Initially, we use 11 different features to implement the model. It is also noticed that not all features are required during the steel processing selection process, and from the feature importance statistics carbon, carbon equivalent, sulfur, phosphorus, yield strength, and elongation are enough to predict the rolling process.



**Figure 7.** Rolling process comparison actual vs predicted results.

The model accuracy metric is often used for determining the performance of various data processing and modeling techniques. In learning models, this metric is often challenged due to the overfitting of the training data for the model. However, when it comes to machine learning, unlike inductive learning, machine learning tries to generalize the solution beyond the training data. Hence, to validate any machine-learning model, especially for greater generalization accuracy cross-validation method is often adopted. This methodology is very effective when training data is limited. Cross-validation splits the given dataset into multiple folds ( $k$  – no of folds) containing both training and testing sets. With the fixation of  $k$ , the model is repeatedly trained on the training set but tested on test sets of all the folds except the trained one. When it comes to testing accuracy, an average of all the accuracy results for all the  $k$ -folds is considered. The dataset is reset to its original state or shuffled to avoid any biasing and then subjected to dividing for validation sets such that proper balance of data distribution can be achieved. In this research, the dataset is divided into 129 samples of training and 10 of testing forming a 10-fold cross-validation set. Figure 8, shows the cross-validation analysis which validates the result of cross-validation datasets which are 80%, 100%,100%,92%,100%, 100%,100%, 100%, 100%,100%, and 90% respectively in each split, and the average prediction

accuracy is thus 96.98 (~97) which proves the acceptability level of prediction of the rolling process selection model.



**Figure 8.** k fold cross validation graph where k=10

#### 4. Conclusion

A computational method has been developed to classify the different process routes of steel constituted by composition and mechanical properties. This study is conducted to develop an effective learning model that can identify thermo-mechanical processing methods for different grades of steels as per targeted mechanical properties. Various outcomes have been studied, including the process of feature selection, data normalization, and the effect of multiple mechanical properties with respect to weight percentage of various elements like carbon, phosphorus, manganese, sulphur, etc present in different grades of steels. This study predicts the steel processing method based on experimental data, and the essence of the same work can be extended to predict properties of different materials by considering their composition and process routes if enough data sets are available. A high and accurate prediction results shown the model has excellent learning and generalization ability. In summary, a simple attempt has been made in the present study to highlight the encouraging convergence among the principles of material science, computing techniques, and informatics, which has been identified as the driver for Industry 4.0. It is reasonable to expect that the approach address in the present study will be extended in future particularly in the domain of materials informatics.

**Ethical Approval:** The author warrant that the paper fulfills the ethical standards of the journal.

**Consent to Participate:** It is confirmed that all the authors are aware and satisfied of the authorship order and correspondence of the paper.

**Consent to Publish:** Author have read and agreed to the published version of the manuscript.

**Authors Contributions:** The work has been done by the corresponding author.

**Funding:** No external funding involved in this research.

**Competing Interests:** The author declare no competing interests

**Availability of data and materials:** NA

### References:

- [1] G. Krauss, Steels: **2015**, Processing, Structure, and Performance, ASM International, Ohio, 217-224.
- [2] W. McFarland, H. Taylor, **1969**, Properties and Applications of Low Carbon Martensitic Steel Sheets, SAE Technical Paper, 690263.
- [3] Bramfitt, B. L. **1998**, Structure/Property relationships in irons and steels. Materials Park, OH: ASM International, 153-173.
- [4] E. Kanca, F. Çavdar and M. M. Erşen, **2016**, Prediction of mechanical properties of cold rolled steel using genetic expression programming, Polish Academy of Sciences, 130, 365-369.
- [5] M. Durand-Charre, **2013**, Microstructure of steels and cast irons, Engineering Materials and Processes, Springer Science & Business Media.
- [6] R. Ueji, N. Tsuji, Y. Minamino and Y Koizumi, **2002**, Effect of rolling reduction on ultrafine grained structure and mechanical properties of low-carbon steel thermomechanically processed from martensite starting structure, Acta Materilia, 50, 4177-4189.
- [7] N. Tsuji, Y. Matsubara, and Y. Saito, **1997**, Dynamic Recrystallization of Ferrite in Interstitial Free Steel, Scripta Materialia, 37, 477-484.
- [8] J. Baczynski, J.J. Jonas, **1998**, Torsion textures produced by dynamic recrystallization in  $\alpha$ -iron and two interstitial-free steels. Metallurgical and Materials Transactions A, 29, 447-462.
- [9] R. Song, D. Ponge, D. Raabe and R. Kaspar, **2005**, Microstructure and crystallographic texture of an ultrafine grained C-Mn steel and their evolution during warm deformation and annealing Acta Materialia 53, 845-858.
- [10] P. D. Hodgson, M. R. Hickson and R. K. Gibbs, **1999**, Ultrafine ferrite in low carbon steel, Scripta Materialia, 40, 00411-4.

- [11] Y. Mazaheri, A. Kermanpur, A. Najafizadeh and N. Saeidi, **2014**, Effects of Initial Microstructure and Thermomechanical Processing parameters on Microstructures and Mechanical Properties of Ultrafine Grained Dual Phase, Steels Material Science Engineering A, 612, 54-62.
- [12] L. Zhao, N. Park, Y. Tian, S. Chen, A. Shibata and N. Tsuji, **2017**, Novel thermomechanical processing methods for achieving ultra-grain refinement of low-carbon steel without heavy plastic deformation Materials, Research. Letters. 5, 61-68.
- [13] J. Zhao and Z. Jiang, **2018**, Thermomechanical processing of advanced high strength steels Progress in. Material. Science, 94, 174-242.
- [14] F. Popa, I. Chicinaş, D. Frunza, I. Nicodim and D. Banabic, **2014**, Influence of high deformation on the microstructure of low-carbon steel, International Journal of Minerals, Metallurgy, and Materials, 21, 273-278.
- [15] A. H. Orta, I. Kayabasi, M. G. Senol, **2019**, Prediction of mechanical properties of cold rolled and continuous annealed steel grades via analytical model integrated neural networks, Ironmaking & Steelmaking.
- [16] L. Ward, A. Agrawal, A. Choudhary and C. Wolverton, **2016**, A general-purpose machine learning framework for predicting properties of inorganic materials, npj Computational Materials, 2, 16028-1.
- [17] Wei Chen, Fanbei Kong, Baoxiang Wang & Yuhan Li, **2019**, Application of grey relational analysis and extreme learning machine method for predicting silicon content of molten iron in blast furnace, Ironmaking & Steelmaking, 46, 974-979.
- [18] A Brahme, M. Winning and D. Raabe, **2009**, Prediction of cold rolling texture of steels using an artificial neural network, Computational Materials Science, 46, 800-804.
- [19] A. González-Marcos, et. al. **2014**, Advanced predictive system using artificial intelligence for cleaning of steel coils, Ironmaking & Steelmaking, 41, 262-269.
- [20] Y. Liu, et. al., 2020, Machine learning in materials genome initiative: A review, Journal of Materials Science & Technology, 57, 113-122
- [21] R. N. Parkins, M. Elices, V. Sanchez-Galvez and L. Caballero, **1982**, sensitive cracking of prestressing steels, Corrosion Science, 22, 379-405.

- [22] A. Choudhury, R. Naskar, A. BasuMallick, S. Pal, 2019, Computer Vision Approach for Phase Identification from Steel Microstructure Journal of Engineering Computations, 36, 1913-1932.
- [23] S. Lalam, P. K. Tiwari, S. Sahoo, A. Dalal, **2019**, Online prediction and monitoring of mechanical properties of industrial galvanised steel coils using neural networks, Ironmaking & Steelmaking, 46, 89-96.
- [24] L. Ward, A. Agrawal, A. Choudhary and C. Wolverton, **2016**, A general-purpose machine learning framework for predicting properties of inorganic materials npj Computational Materials, 2, 16028.
- [25] H. Pouraliakbar, M. J. Khalaj, M. Nazerfakhari and G. Khalaj, **2015**, Artificial neural networks for hardness prediction of HAZ with chemical composition and tensile test of X70 pipeline steels, Journal of Iron and Steel Research International, 22,, 446-450.
- [26] X. Li, Y. Fang, L. Liu, Kernel, **2020**, extreme learning machine for fatness pattern recognition in cold rolling mill based on particle swarm optimization, Journal of the Brazilian Society of Mechanical Sciences and Engineering, 42, 270
- [27] L. Zheng, Y. Xiang, C. Sheng, **2019**, Optimization-based improved kernel extreme learning machine for rolling bearing fault diagnosis, Journal of the Brazilian Society of Mechanical Sciences and Engineering, 41, 505.
- [28] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi and C. Kim, **2017**, Machine Learning and Materials Informatics: Recent Applications and Prospects npj Computational Materials, 3, 54.
- [29] B. Medasani and I. Vasiliev, **2009**, Computational study of the surface properties of aluminum nanoparticles , Surface Science, 603, 2042-2046.
- [30] S. Dwarapudi, P. K. Gupta & S. S. Gupta, **2006**, Application of artificial neural network model to predict reduction degradation index of iron oxide pellets, Ironmaking & Steelmaking, 33, 500-506.
- [31] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel and J. Vanderplas, **2011**, Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 2825-2830.
- [32] K P Murphy, Machine Learning: **2012**, A Probabilistic Perspective, MIT press.

- [33] B. L. Bramfitt, **1998**, Structure/Property relationships in irons and steels. Materials Park, ASM International, 153-173.
- [34] A. Choudhury, T. Konnur, P.P. Chattopadhyay, S. Pal, **2019**, Structure prediction of multi-principal element alloys using ensemble learning, *Engineering Computations*, 37 No. 3, 1003-1022.
- [35] G. Rossum, **1995**, Python Reference Manual, Technical Repository, Amsterdam.
- [36] D. M. Farid, M. M. Rahman, M. A. Al-Mamuny, **2014**, Efficient and scalable multi-class classification using naïve Bayes tree, *International Conference on Informatics, Electronics & Vision*, 1-4.
- [37] M. Brammer, **2010**, Technology and its community in today's metals industries, *Ironmaking & Steelmaking*, 37, 265-270.
- [38] C. M. Bishop, **2006**, Pattern recognition and machine learning, Springer.
- [39] T.K. Ho, **1995**, Random decision forests in document analysis and recognition, *International Conference on Document Analysis and Recognition IEEE*, 1, 278–282.
- [40] Demir, H. (2018). Classification of texture images based on the histogram of oriented gradients using support vector machines. *Istanbul University Journal of Electrical & Electronics Engineering*, 18(1), 90–94.
- [41] Joshi, K. D., Chauhan, V., & Surgenor, B. (2018). A flexible machine vision system for small part inspection based on a hybrid SVM/ANN approach. *Journal of Intelligent Manufacturing*
- [42] S. Pathak, I. Mishra and A. Swetapadma, **2018**, "An Assessment of Decision Tree based Classification and Regression Algorithms," *2018 3rd International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, 92.
- [43] S. Koley, T. Ray, I. Mohanty, S. Chatterjee, M. Shome, **2019**, Prediction of electrical resistivity of steel using artificial neural network, *Ironmaking & Steelmaking*, 46, 383-391.

# Figures

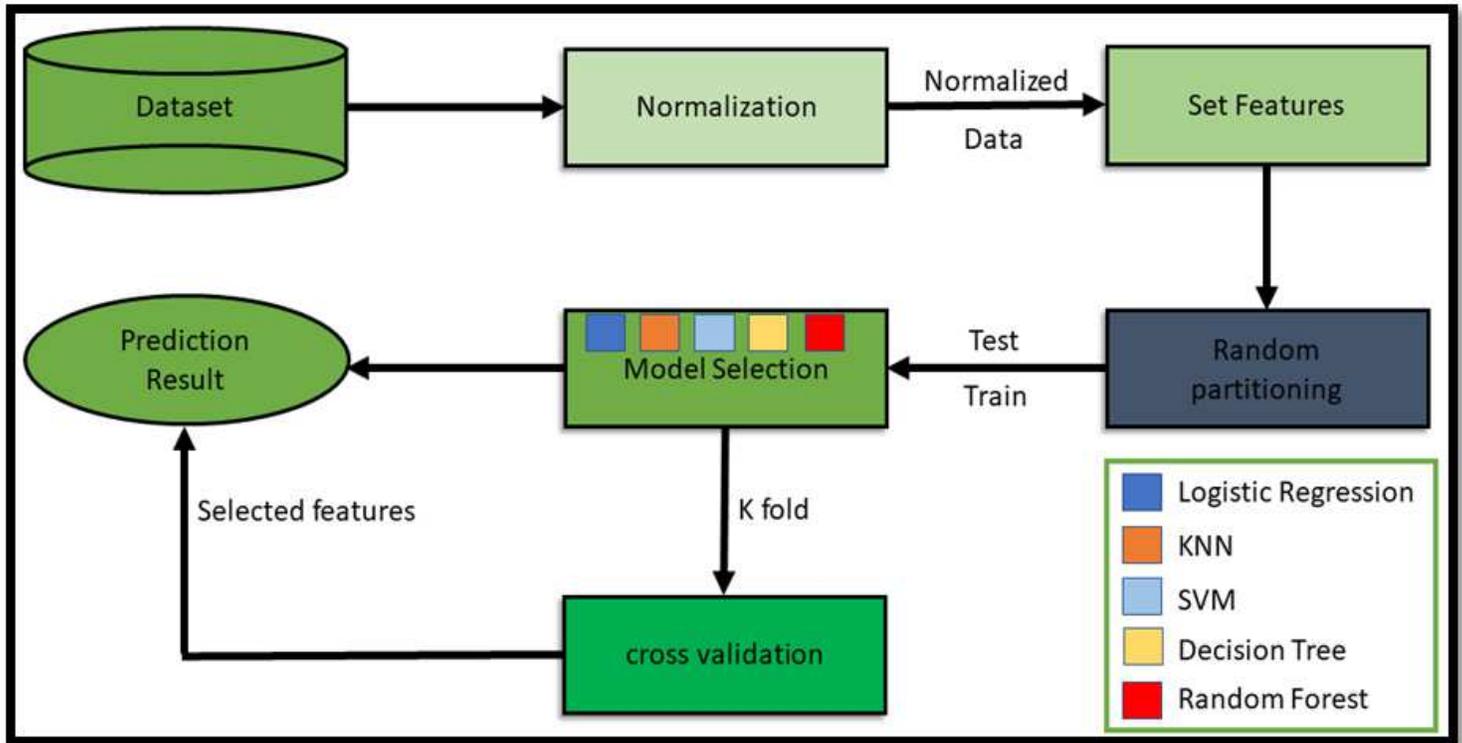


Figure 1

Schematic representation of steel processing predictive model.

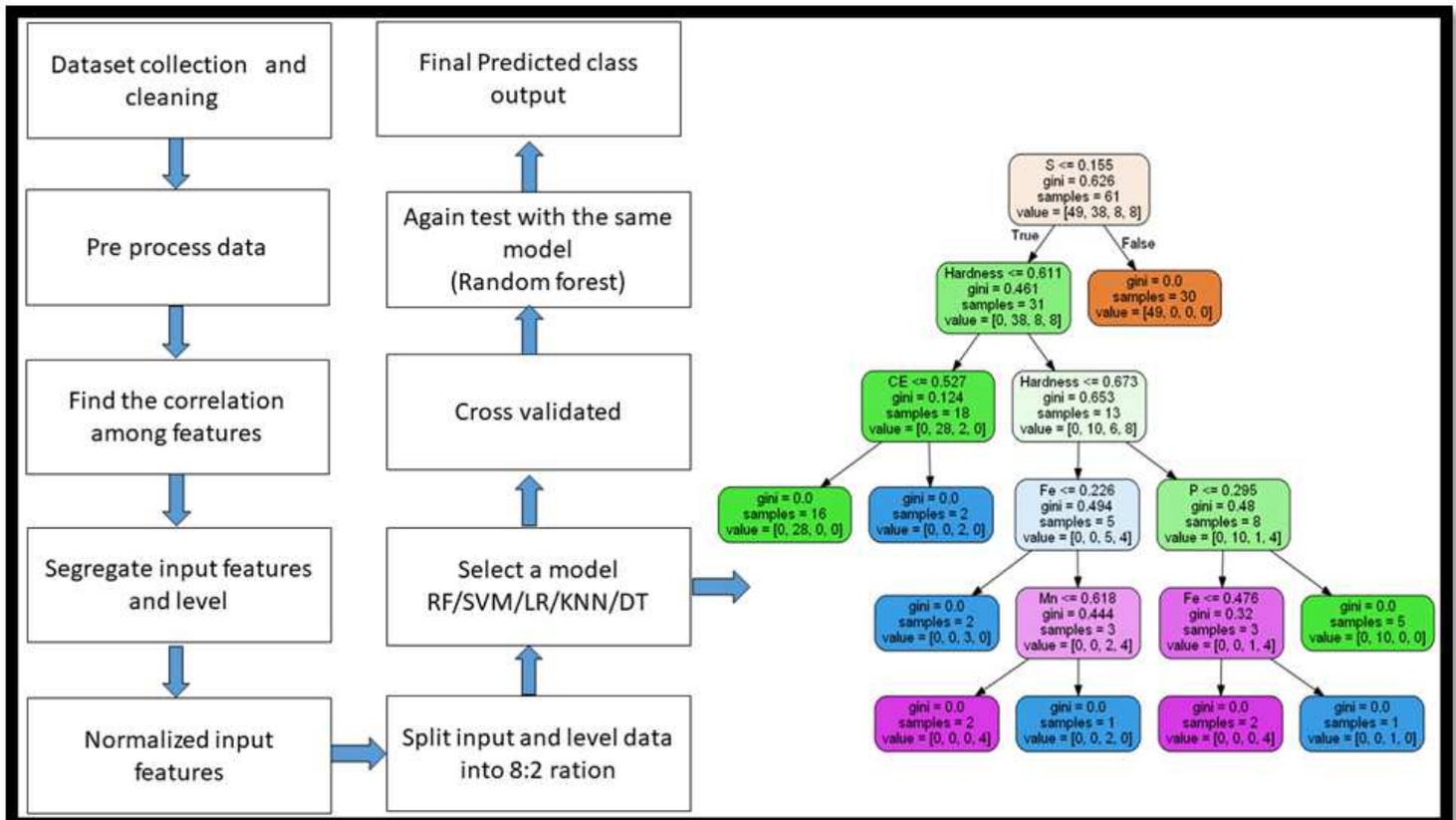


Figure 2

Process steps involved for the model development. The tree in the right side shows how data slit into the left and right node so that optimal results can be derived by using a voting scheme.

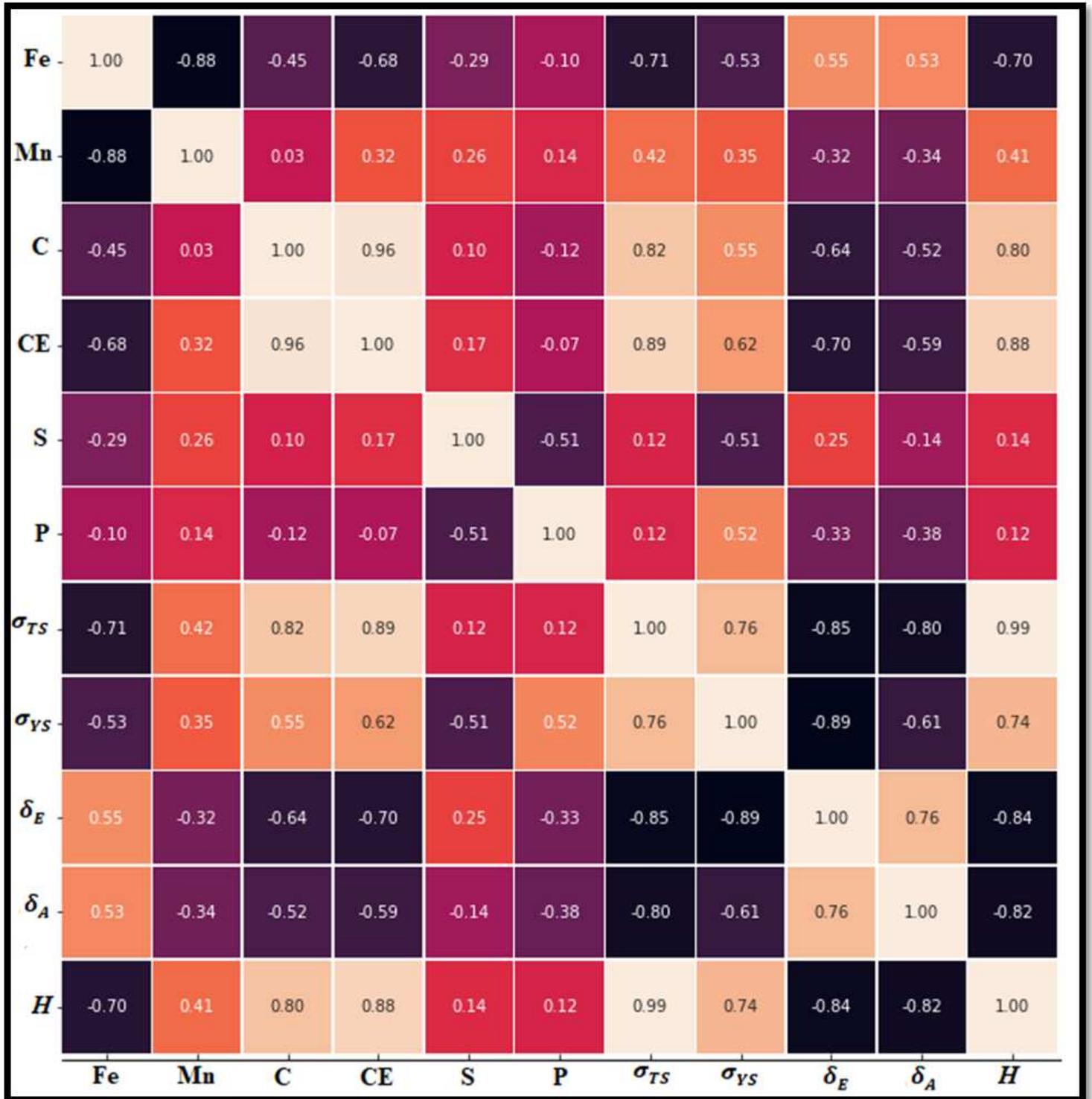


Figure 3

Correlation map: X and Y-axis parameters are Iron, Manganese, Carbon, Carbon equivalent, Sulfur, Phosphorous, Tensile Strength, Yield strength, Elongation, Reduction in area, Hardness.

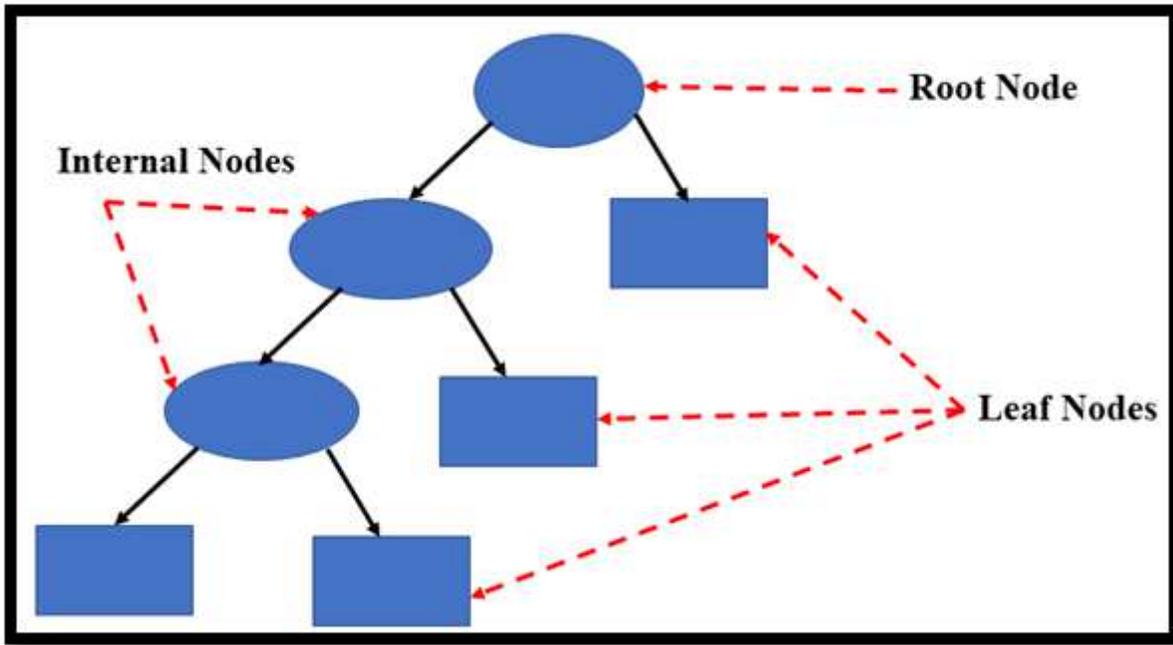


Figure 4

Node block diagram of a decision tree. Left-hand side nodes are known as the left child node, and the right side nodes are known as the right child node.

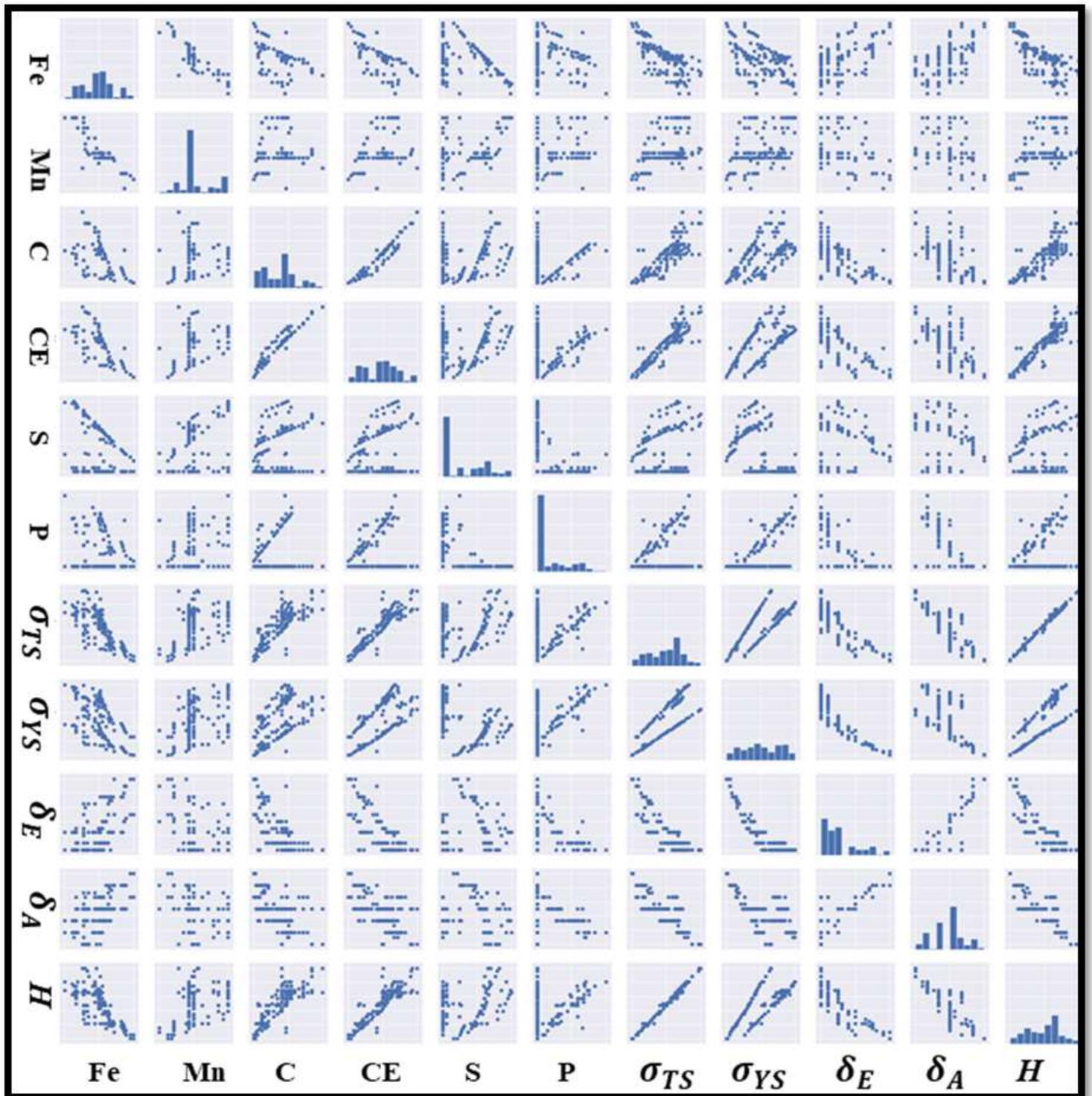


Figure 5

Correlation among various features using a scatter plot. The diagonal matrix indicates the highest correlation, among other features.

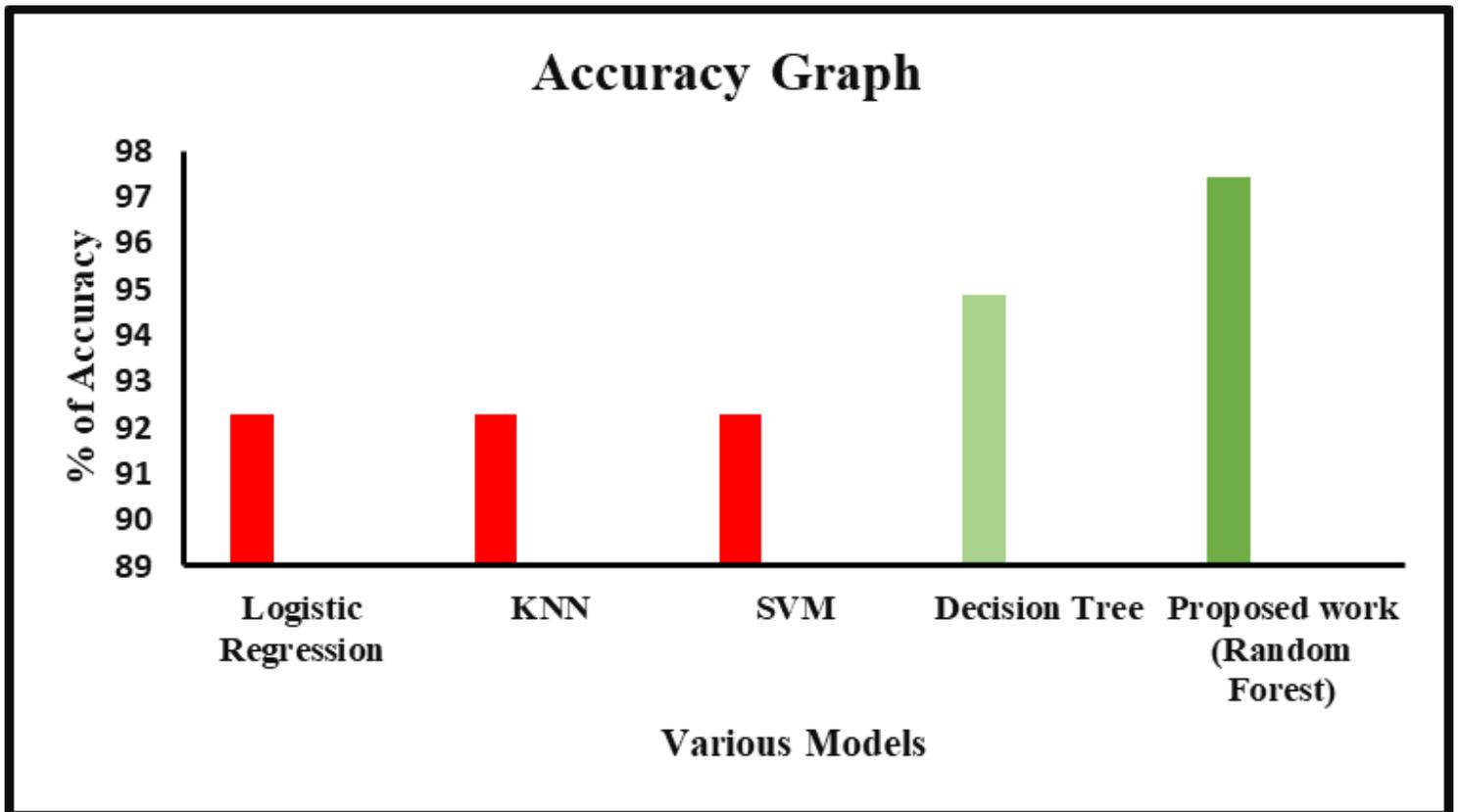


Figure 6

Comparisons of the accuracy of different machine learning models.

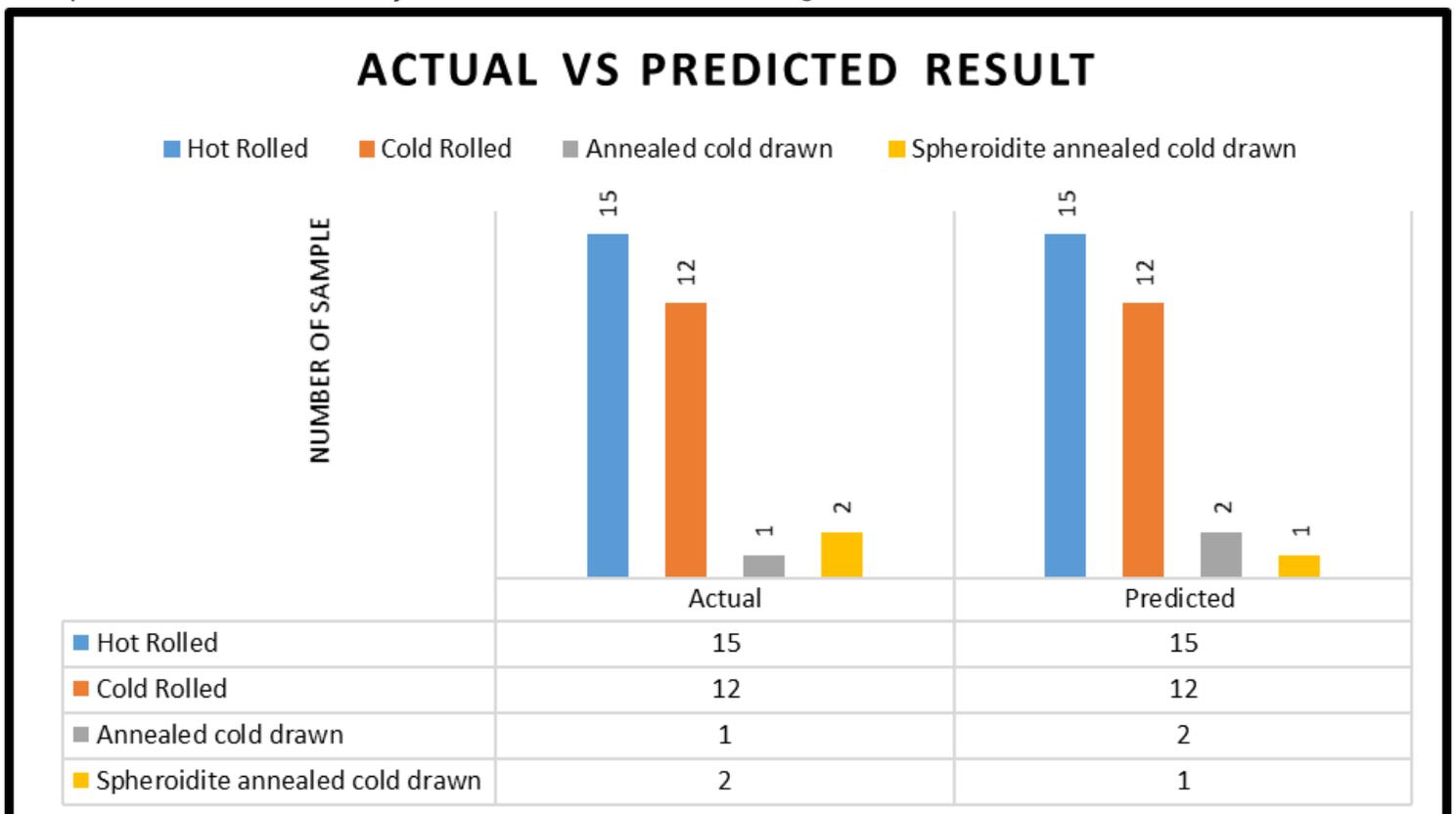


Figure 7

Rolling process comparison actual vs predicted results.

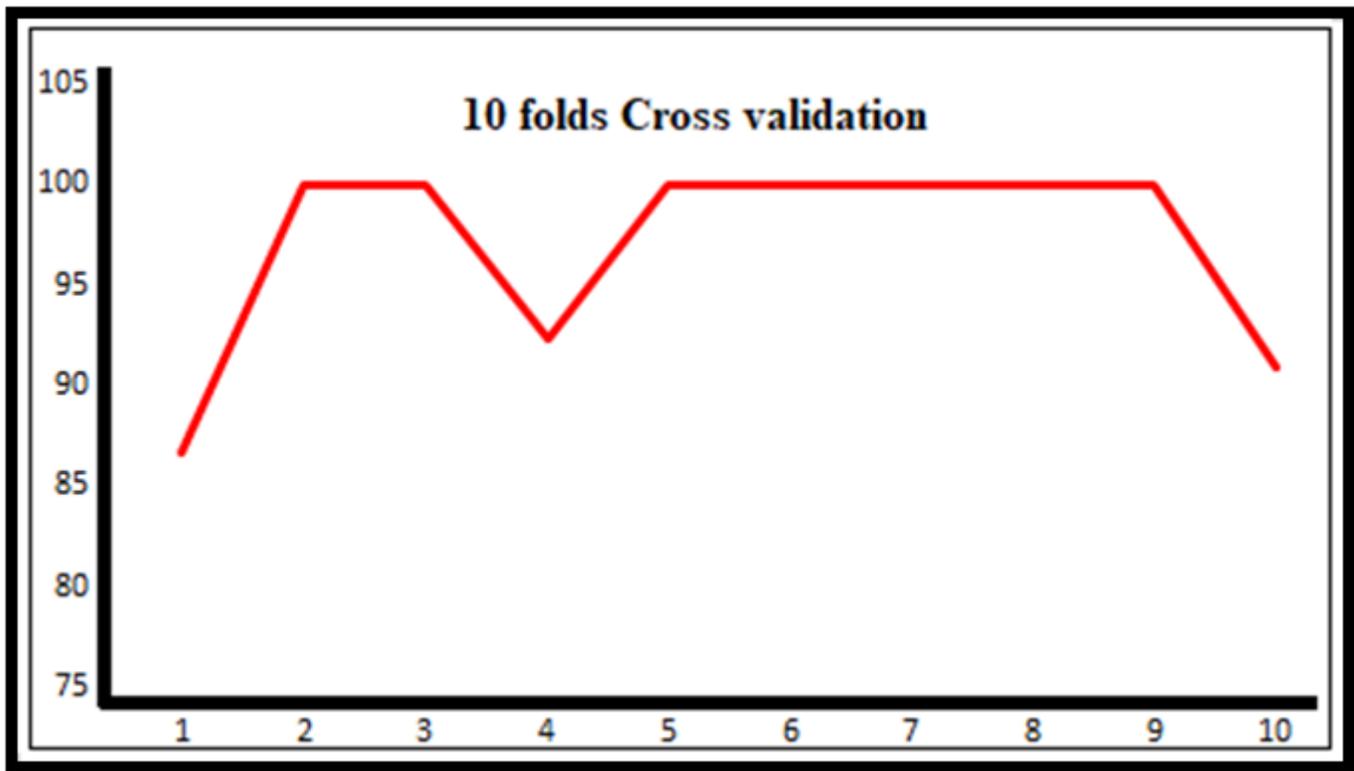


Figure 8

k fold cross validation graph where k=10