MACHINE LEARNING MODELS FOR PREDICTING MECHANICAL PROPERTIES IN FRICTION STIR WELDING OF AL ALLOYS

B. Gugulothu[1], K. Srividya[2], S. Vijayakumar[3,*]
I. Veeranjaneyulu[4], S. Revathi[5], M. Ramya[6]

[1]Saudi Electrical Services Polytechnic, Ras Tanura, Saudi Arabia
[2]Department of Mechanical Engineering, P V P Siddhartha Institute of Technology, Kanuru, India
[3]Department of Mechanical Engineering, Saveetha School of Engineering, SIMATS, Saveetha University, Tamil Nadu, India
[4]Department of Mechanical Engineering, Aditya University, Surampalem, India
[5]Department of Research and Innovation, Saveetha School of Engineering, SIMATS, Saveetha University, Tamil Nadu, India
[6]Division of Research and Development, Lovely Professional University, Phagwara, Punjab, India
*Corresponding author's e-mail address: vijaysundarbe@gmail.com

ABSTRACT

Friction Stir Welding (FSW) has developed as an extremely reliable solid-state joining technique for Al alloys, which possesses superior mechanical properties and minimal defects compared to conventional fusion welding. Accurate prediction of welding output parameters such as Ultimate Tensile Strength (UTS), Elongation (%), Hardness, and Wear Rate is crucial for ensuring weld quality and optimizing process conditions. In this study, four machine learning (ML) approaches were employed, such as Backpropagation Neural Network (BP), Extreme Gradient Boosting (XGB), Support Vector Machine (SVM), and Principal Component Analysis (PCA) to predict the output responses with the help of a dataset derived from experimental values. Initial correlation analysis highlighted D/d ratio and tilt angle as critical parameters, moderately correlated with UTS, elongation, and wear rate, but weld speed was inversely related to hardness. In the model performance, SVM achieved the highest prediction accuracy of more than 99.5% for UTS and elongation, followed by BP and XGB methods. PCA demonstrated stable performance, particularly for hardness prediction and BP achieved ~99.44% accuracy for wear rate. Error analysis demonstrated that SVM exhibited the lowest and most stable percentage errors, particularly for UTS and elongation, while PCA showed the least deviation for hardness predictions. The Mean Squared Error (MSE) values for the models were as follows: BP 0.6280, XGB 1.1681, SVM 0.4838, and PCA 0.9590. Overall, the comparative study validates the effectiveness of integrating ML algorithms for accurate welding output prediction, offering potential for real-time process monitoring and optimization in advanced manufacturing environments.

KEYWORDS: machine learning, welding, BP neural network, XGBoost, SVM, PCA, prediction, mechanical properties, Friction Stir Welding

1. INTRODUCTION

FSW is a solid-state joining technique that has gained significant attention in the manufacturing industry, particularly for welding aluminium alloys and other materials that are challenging to join using traditional fusion welding methods [1-2]. This technique is widely adopted in the automotive and aerospace industries due to its several advantages over the conventional welding process [3]. This process involved a non-consumable rotating tool that generates frictional heat and plastic deformation as it moves along the joint line, creating a weld without melting the base materials [4]. This solid-state nature of FSW results in several benefits, including reduced heat-affected zone (HAZ), excellent mechanical and corrosion resistance properties, and the elimination of unfavourable phase transformations that can occur in fusion welding [5]. Generally, parameters such as tool rotational speed, welding speed, tilt angle, D/d ratio (shoulder diameter to pin diameter ratio), and tool geometry in FSW play crucial roles in determining

Date of submission/acceptance: 06.08.2025/15.12.2025
https://doi.org/10.35219/awet.2025.15

the quality and strength of the welded joints. Welding speed and tool rotational speed were the most significant factors affecting the weld strength [6-7]. Increasing rotational speeds leads to increased heat generation and better material mixing, while appropriate welding speeds ensure proper material flow and consolidation, and the D/d ratio affects the heat input and material flow during the welding process. A larger shoulder diameter typically results in increased heat generation and a wider heat-affected zone [8]. The optimization of FSW for aluminium alloys continues to be an active area of research, with studies focusing on various aspects such as process parameters, tool design, and material combinations. The integration of machine learning and AI techniques offers new avenues for further optimization and quality control in FSW of aluminium alloys, potentially increasing the acceptance and reliability of the process in high-integrity applications.

The different optimization approaches have been employed to improve FSW performance. For instance, Tamjidy et al. [9] proposed a multi-objective algorithm based on biogeography-based optimization to enhance mechanical properties in the heat-affected zone. This approach utilizes decision-making techniques like Shannon's entropy and TOPSIS to select the optimal solution. Abbasi et al. [10] employed FEM simulation to evaluate temperature distribution during FSW, enabling the characterization of microstructure in different zones and optimizing welding variables with reduced cost and energy concerns. Prabhakar et al. [11] highlighted the integration of signal processing with artificial intelligence and machine learning algorithms to improve weld quality. These varied approaches underscore the complexity of FSW optimization and the need for tailored solutions based on specific materials and applications.

Future research should focus on combining multiple optimization techniques to achieve comprehensive improvements in FSW performance across different materials and joint configurations. In recent years, the utilisation of machine learning has significantly expanded in addressing the intricate dynamics of manufacturing processes. These trained algorithms can rapidly assimilate historical data and elucidate the intricate interconnections among diverse processing parameters [12]. Many researchers have explored the use of ANNs [13], ANFIS [14], and regression analysis [15] to prioritize, optimize, and forecast experimental results. Additionally, deep neural networks have shown promise in evaluating weld quality and detecting welding defects based on recorded data. RNN and CNN were trained to classify FSW force data sets, achieving detection accuracies over 95% for weld defects larger than 0.08 mm [16]. Kumaresan et. al [17] studied Image Processing Techniques (IPTs) for detecting and classifying weld defects. A total of 940 image patches representing various weld defects were sourced from the GDXray database and systematically categorized and they adopted DNNs combined with transfer learning approaches. Among the tested models, the Support Vector Machine (SVM) classifier using features derived from ResNet50 achieved the highest performance, reaching an impressive accuracy of 98%. Overall, the proposed models demonstrated superior test accuracies of 99.4% for nine-class defect classification and 97.8% for fourteen-class classification.

Distun Stephen et al. [18] employed deep learning techniques to autonomously detect different categories of weld flaws through radiographic pictures. They develop and train a CNN model utilising 200 annotated radiographic pictures, each demonstrating one of four categories of welding faults. Firstly, the model demonstrated overfitting without data augmentation, yielding an accuracy of 80% and a loss of 6.7. But the CNN model achieves 95% validation accuracy and 99.2% training accuracy.

Another study developed an ML framework to forecast the tensile strength of FSW-produced aluminium alloys using a dataset of 213 samples. Various ML algorithms were applied, with the adaptive boosting classifier (ABC) achieving the highest accuracy of 81.6% [19]. Another research optimized FSW parameters for 2024T3 aluminium alloy sheets using ML models, including random forest, XGBoost, and multilayer perceptron artificial neural network (MLP-ANN). These models achieved over 98% accuracy in parameter regression, leading to an FSW joint efficiency of 93% relative to the base material [20]. The use of ANN and SVM in optimizing the FSW process, highlighting their strong predictive capabilities and efficiency in process optimization [21]. Sarsılmaz et.al [22] reported that ANN models outperformed SVM techniques in fine-tuning FSW parameters, leading to significant improvements in the tensile strength and hardness of the welded joints.

Jianming Hu et al. [29] analysed RSW process signals of 2219/5A06 aluminium alloy under two assembly conditions – gap and spacing – and applied artificial intelligence modelling to improve the accuracy and efficiency of RSW quality evaluation. A multi-signal fusion approach was proposed, combining principal component analysis and correlation analysis. To optimize the model, a backpropagation neural network (BPNN) was enhanced using the sine-chaotic-map-improved sparrow search algorithm (SSA). The model's input and output consisted of the variables resulting from multi-signal fusion and the button diameter, respectively.

In another study [30], laser–arc hybrid welding was used to perform butt welding on 6 mm Q345 steel under various assembly conditions. An adaptive model of the BP neural network, optimized by a genetic algorithm (GA), was proposed for this laser–arc welding process. By optimizing the neural network parameters using the GA, a relationship between the pre-welding groove parameters and welding

parameters was established. The GA-BP neural network showed a mean square error (MSE) of just 0.75%. Experimental verification confirmed that the neural network could predict the welding parameters required to achieve specific welding morphologies based on different pre-welding grooves. Additionally, a four-dimensional mapping relationship was developed between welding parameters and the mechanical properties of joints through the analysis of friction stir welding (FSW) data. This enabled intelligent optimization of the welding process and the prediction of joint properties. The BPNN model revealed the metallurgical mechanisms behind the mapping process-property relationship, and the proposed method achieved a prediction accuracy of 92%. The welding parameters optimized by the BPNN model were 1810 rpm for rotational speed, 105 mm/min for welding speed, and 3 kN for welding pressure. These conditions resulted in a joint tensile strength of 415 MPa, with a deviation of only 3.71% from the experimental value [31].

Sadra Naddaf-Sh et al. [32] applied CSSA-XGBoost algorithm was to predict the tensile shear strength of resistance spot welding joints by extracting representative features from dynamic resistance signals during the welding process. The method involved using an improved sparrow search algorithm to optimize XGBoost hyperparameters, incorporating a chaotic map and Levy flight strategy to enhance the quality of the hyperparameter search. Comprehensive experiments conducted on the self-built dataset demonstrated that the prediction achieved a root mean square error (RMSE) of 0.3423 KN and a mean absolute percentage error (MAPE) of only 2.25%. The proposed approach outperformed other algorithms, providing a reliable method for weld quality inspection in industrial applications.

Based on the literature reviews, several machine learning approaches have been applied to predict welding quality and associated mechanical properties. However, existing studies utilized no more than three techniques and performed comprehensive comparisons across multiple models.

- Addressing this gap, the present work introduced a novel approach by employing four distinct machine learning methods: BP Neural Network, XGB, SVM, and PCA to predict output responses.
- These models were trained and evaluated using a dataset derived from experimental measurements of welded joints in aluminium alloy, offering a more extensive comparative analysis.
- This work also analyses the actual versus predicted values of output responses, providing insights into the models' ability to predict welding quality accurately. Performance metrics such as MAE, MSE, and RMSE were used to evaluate the models.
- Furthermore, the comparison of error percentages for output responses highlights the relative performance of each model, offering a robust evaluation of their prediction accuracy.

2. DATASET COLLECTION

Data collection is a critical preliminary step in the development of any machine learning model. In this study, data were sourced primarily from source [23]. According to data collection [23], Al-Mg-Mn alloy measuring 160mm × 100 mm × 5 mm was selected as the workpiece material for experimentation and validation. Friction Stir Welding (FSW) was carried out using a tapered cylindrical tool [33]. A schematic illustration of the FSW setup for dissimilar joint formation is shown in figure 1. Four critical process parameters were identified for the optimization of the FSW process: rotational speed (RD), welding speed (WS), tilt angle (TE), and shoulder-to-pin diameter ratio (D/d). The experimental ranges for these input variables were established as follows:

- RD: 1000-1400 rpm,
- WS: 80-120 mm/min,
- TE: 1.5 -2.5,
- D/d Ratio: 2.25-3.8.

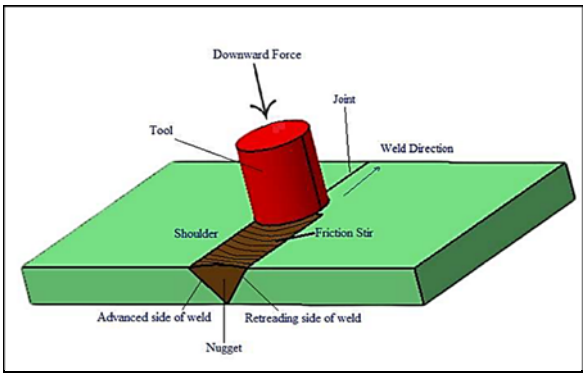

Fig. 1. Schematic diagram of the FSW process

Experiments were designed and executed according to the BBD matrix. Following the welding trials, four key tests were conducted on the samples to assess performance: Ultimate Tensile Strength (UTS), Elongation (%), Brinell Hardness Number (HBN), and Wear Rate (WTR). From the experimental output data, only 15 set of results was selected for predictive analysis. These outputs were modelled using four machine learning techniques: Backpropagation Neural Network (BP), Extreme Gradient Boosting (XGB), Support Vector Machine (SVM), and Principal Component Analysis (PCA) to evaluate and compare their prediction accuracy. The flow of this study is shown in figure 2.

3. METHODOLOGY

Four models were trained using standardized inputs, with output responses predicted and compared against actual values.

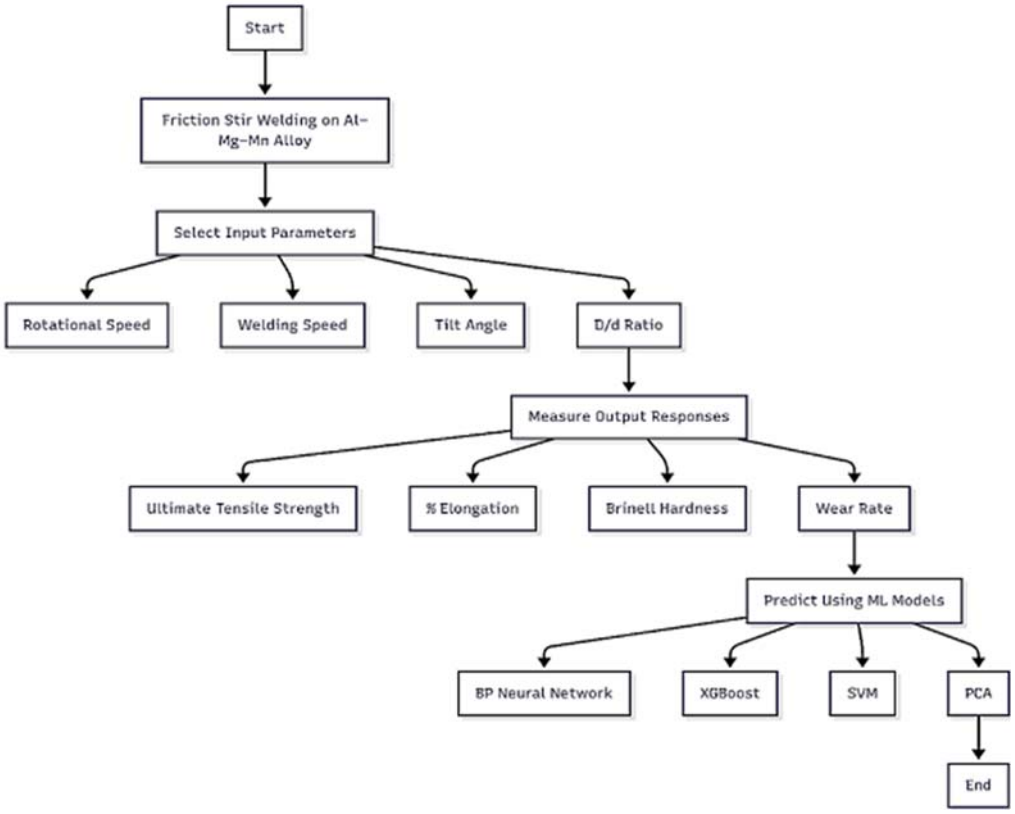

Fig. 2. The flow process of present work

3.1. BP Neural Network (BPNN)

BPNN architecture is particularly effective for tasks involving nonlinear mapping, such as pattern recognition and function approximation, due to its ability to learn complex relationships within data. This model typically consists of three layers: an input layer, hidden layers, and an output layer. the input layer was used to find the type and volume of incoming data.

The hidden layer introduces nonlinearity into the model through its structure, determined by the number of neurons and activation functions, enabling the network to capture intricate patterns. Finally, the output layer was responsible for generating the final predicted values based on the processed inputs [24]. The output of the neuron model structure is usually expressed in Equation (1)

$$\hat{x}_j = e\left(\sum_{j=1}^{m} w_{ji}x_j - b_{ji}\right) \quad (1)$$

Typical loss function for optimization is shown in Equation (2).

$$E = \sum_{j=1}^{m} \left(\hat{x}_j - x_j\right)^2 \quad (2)$$

where $\widehat{xj}$ is the output of the neuron, aji;
vji - the bias weights of the neuron;
E is a loss function.

Weight minimization is calculated by equation (3)

$$v_{ji} = v_{ji} - \eta \frac{\partial E(v_{ji})}{\partial v_{ji}} \quad (3)$$

3.2. XG Boost

It is a powerful ensemble learning method built on the foundation of gradient-boosted decision trees and has gained widespread adoption across diverse application domains. The central concept behind XGBoost is to combine multiple weak learners, typically decision trees, through the gradient boosting framework.

In each iteration, a new tree is added to model the residual errors from the previous iteration. The formulation of this objective function is shown in equation (4)

$$L(f) = \sum_j l\left(\hat{x}_j, x_j\right) + \sum_l \Omega(f_l) \quad \text{where } \Omega(f) = \gamma T + \frac{1}{2}\lambda\|\omega\|^2 \quad (4)$$

$l(\widehat{x}j,xj)$ - the loss function;
$\hat{x}j$ - the predicted value and xj is the actual value;
$\Omega(fl)$ - the regularization term, which is used to control the model complexity;
fl - the model of the l-th tree;
T - the number of leaf nodes in the l-th tree;
ω - the weight of the leaf nodes in the l-th tree;
Y - the penalty regularization term for the leaf nodes;
λ - the penalty regularization term for the leaf weights [25].

3.3. SVM Algorithm

It follows a systematic approach to classify or predict outcomes based on input features. Initially, labelled training data consisting of input-output pairs is provided. The algorithm then selects a suitable kernel function to transform the data, especially if it's not linearly separable. During the training phase, SVM identifies the optimal hyperplane that maximizes the margin between classes by focusing on the closest data points. For regression tasks, it fits a line or curve within a margin of tolerance. Once trained, the model uses a regression function to predict the output shown in equation (5):

$$f(\mathrm{y}) = \mathrm{sign}\left(\sum_{j=1}^{m} \alpha_j x_j K(\mathrm{y}_j, \mathrm{y}) + b\right) \quad (5)$$

where:
y is new input vector (test data) to be classified;
f(y) - output prediction: usually +1 or -1;
sign(.) - sign function: returns +1 if input > 0, else-1;
m - number of training data points;
α_j - learned weight (Large multiplier) for training sample j;
x_j - class table of training point j, either +1 or -1;
y_j - training data point j;
K (y_j,y) - Kernal function: measures similarity between y_j and new input y;
b - bias term from training.

3.4. Principal Component Regression (PCR)

It is a widely used statistical technique that combines Principal Component Analysis (PCA) for dimensionality reduction with regression methods for prediction. It is particularly effective in high-dimensional datasets or when multicollinearity is present, as it reduces the number of features while retaining most of the variability in the data. PCR enables more robust and interpretable models by transforming the feature space via PCA before applying a regression model.

The general predictive model using PCR is given by equation (6)

$$x = (y - \mu) \cdot \mathrm{W} \cdot \beta \quad (6)$$

where β is regression coefficient;
μ - mean vector of training input.

3.5. Performance metrics

The implementation of all models was programmed in Python using machine learning. To validate the performance of the optimized models, six key indicators were introduced, including the coefficient of determination ($R2$), mean absolute error (MAE), mean absolute percentage error (MAPE), mean squared error (MSE), root mean squared error (RMSE), and accuracy [28] [20]. The formula for Performance metrics is shown in equations (7) to (12).

$$R^2 = 1 - \frac{\sum_{i=1}^{m} (\mathrm{x} - \hat{x}_i)^2}{\sum_{j=1}^{m} \left(x_j - \overline{x_j}\right)^2} \quad (7)$$

$$MAE = \frac{1}{m}\sum_{j=1}^{m} |\mathrm{x} - \hat{x}_i| \quad (8)$$

$$MAPE = \frac{100\%}{m}\sum_{j=1}^{m} \left|\frac{x_j - \hat{x}_i}{x_j}\right| \quad (9)$$

$$MSE = \frac{1}{m}\sum_{j=1}^{m} \left(x_j - \hat{x}_i\right)^2 \quad (10)$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{j=1}^{m} \left(x_j - \hat{x}_i\right)^2} \quad (11)$$

$$\mathrm{ACCURACY} = 100\% - \mathrm{MAPE} \quad (12)$$

where m is the number of observations (or data points);
xj - the actual value, $\hat{x}$;
j - the predicted value by the model;
xi - the average of all actual values x.

Furthermore, for each model, comparison plots of the model's predictions against the actual values and learning curve plots were created.

4. RESULTS AND DISCUSSION

The heatmap illustrates the Pearson correlation coefficients between the input parameters and output responses in the welding process. Among the findings, the D/d ratio exhibits a moderate positive correlation with both ultimate tensile strength (0.35) and wear rate (0.34), indicating its significant influence on mechanical performance and material degradation (Fig. 3). Similarly, the tilt angle shows a positive relationship with elongation percentage (0.34) and wear rate (0.28), suggesting that tool inclination affects ductility and surface interaction during welding.

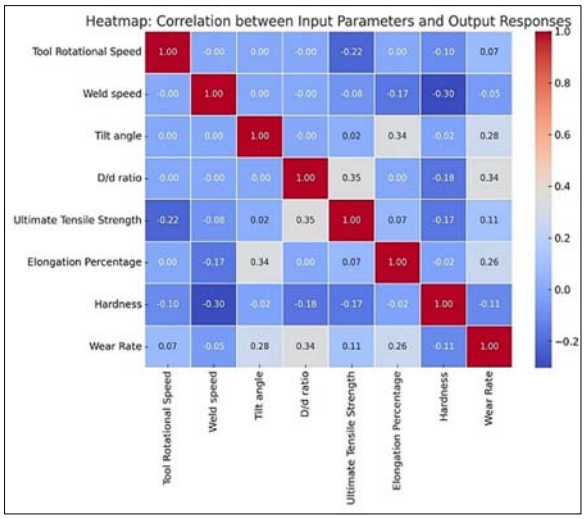

Fig. 3. Heat map: correlation between input parameters and output responses

Welding speed is moderately negatively correlated with hardness (−0.30), implying that higher welding speeds may lead to softer weld zones. Tool rotational speed demonstrates a weak negative correlation with tensile strength (−0.22), but has minimal influence on other outputs. Notably, the input parameters exhibit low intercorrelation among themselves (mostly near zero), which reduces multicollinearity and promotes the application of multivariate techniques such as principal component analysis (PCA) or multiple regression. Overall, the correlation analysis highlights the D/d ratio, tilt angle, and weld speed as the most impactful parameters on the final weld quality and wear characteristics.

4.1. Analysis of Actual versus Predicted Values

The scatter plot provided a comparison between actual UTS and predicted UTS values generated by four different models: BP, XGB, SVM, and PCA.

From figure 4, the diagonal line represents the ideal prediction line, where the predicted values exactly match the actual UTS values. Data points for all four models are tightly clustered around this line, indicating high prediction accuracy across models. Among the four models, BP (blue dots) and XGB (green dots) exhibited near-perfect alignment with the ideal line, which indicated the minimal error. SVM (red dots) and PCA (purple dots) also performed well but show slightly more deviation at higher UTS values, indicating marginally less accuracy.

Scientifically, this performance disparity is attributed to model sensitivity. BP and XGB models were capable of capturing complex nonlinear relationships more effectively due to their deep learning and ensemble-based architectures respectively. Overall, all models demonstrate robust predictive performance, but BP and XGB models slightly outperformed others in capturing the nonlinearities present in the welding process data.

From figure 5, The scatter plot presents a comparison between the actual and predicted values of elongation percentage using four predictive models: from this graph, Data points corresponding to all models are closely aligned with this ideal line, indicating a high degree of prediction accuracy across the models. But BP and PCA predictions are most tightly clustered along the diagonal, suggesting high model reliability and generalization performance for elongation prediction. XGB and SVM predictions also closely follow the ideal trend but show slight deviation in the mid-range values (around 11–13%), indicating potential local overfitting or underfitting effects.

Figure 6 shows the comparison between actual and predicted hardness values generated by four different models. Among the models, PCA and SVM predictions showed the closest adherence to the ideal line across the full range of hardness values (from 94 to 126), with minimal scatter, which suggested the consistent predictive reliability, especially in medium-to-high hardness regions. XGB and BP were with slight dispersion in the lower range of hardness values (around 95–100), possibly due to their sensitivity to local variance or overfitting tendencies when trained on smaller datasets.

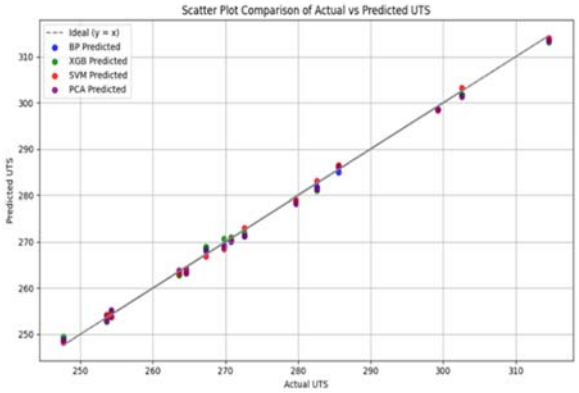

Fig. 4. Scatter plot of actual vs. predicted values for UTS

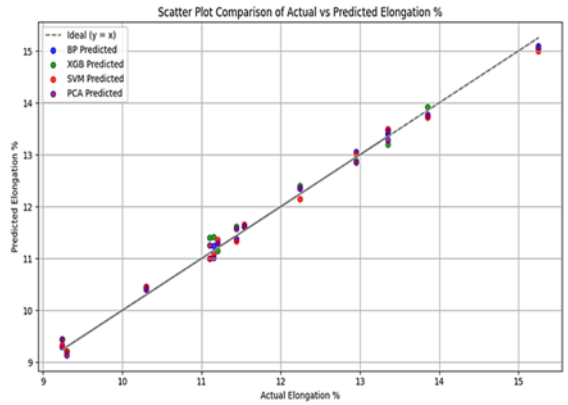

Fig. 5. Scatter plot of actual vs. predicted values for % elongation

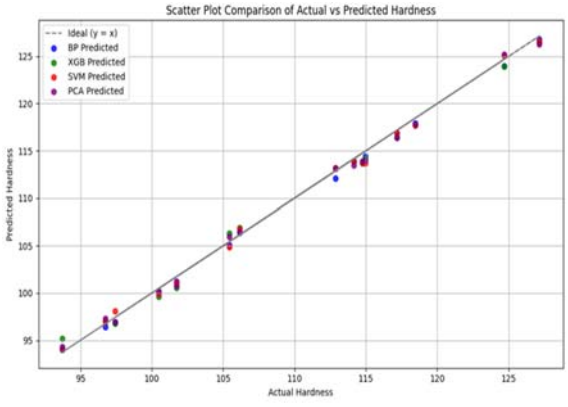

Fig. 6. Scatter plot of actual vs predicted values for hardness

From figure 7, the scatter plot compares the actual and predicted wear rate values using four different models. Among the four models, BP and PCA models

exhibited slightly better alignment with the actual values, particularly at the lower and upper ends of the wear rate spectrum, suggested more consistent performance. XGB shows minor underestimation in higher ranges, and SVM demonstrated reliable performance with slight variations. This analysis highlighted the capability of these models to capture complex, nonlinear relationships inherent in wear rate data, which is crucial in applications such as tool condition monitoring, material degradation assessment, and predictive maintenance.

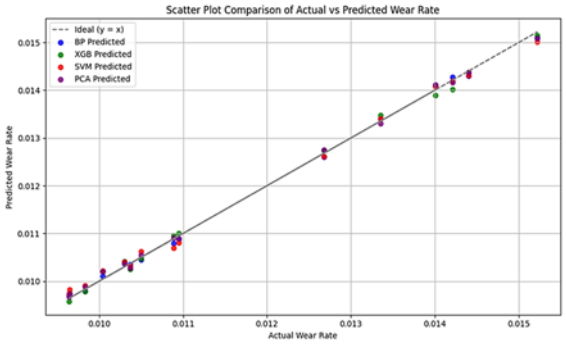

Fig. 7. Scatter plot of actual vs predicted values for wear rate

4.2. Comparison of Error Percentage

The bar chart presents a comparative percentage error analysis on UTS for four predictive models across 15 samples, as figure 8 shows. Across all samples, the highest percentage error was detected in XGB-Sample 2, reaching +0.7% which indicated the greatest overestimation. In contrast, the major negative error appears in XGB-Sample 3 (–0.5%), making XGB the most variable model in terms of error magnitude. BP also showed noticeable deviations, particularly in samples 1, 2, and 6, but with slightly smaller error margins than XGB. SVM and PCA exhibited lower and more consistent error percentages across most samples, with SVM showing a maximum positive error of around +0.25% (Sample 4) and PCA peaking at about +0.4% (Sample 15). Overall, SVM and PCA demonstrate the least percentage error variation, suggesting higher prediction stability, while XGB and BP showed more fluctuations and contributed both the extreme positive and negative deviations. The bar chart displayed the percentage change in prediction error in % of elongation for four models across 15 samples, offering insight into each model's relative performance (Fig. 9).

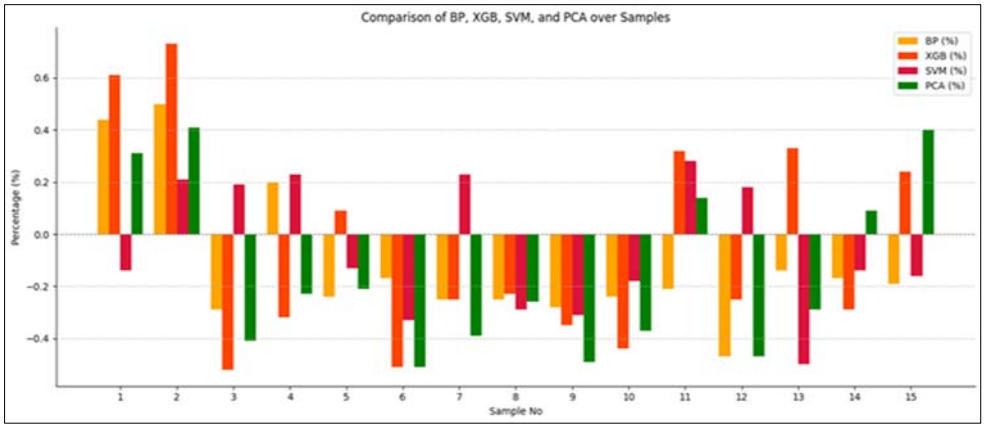

Fig. 8. Comparison plot on UTS

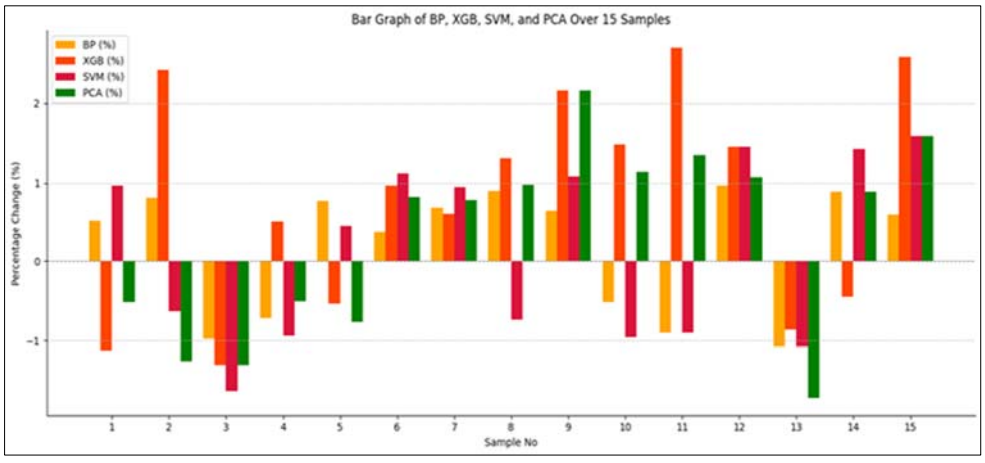

Fig. 9. Comparison plot on % elongation

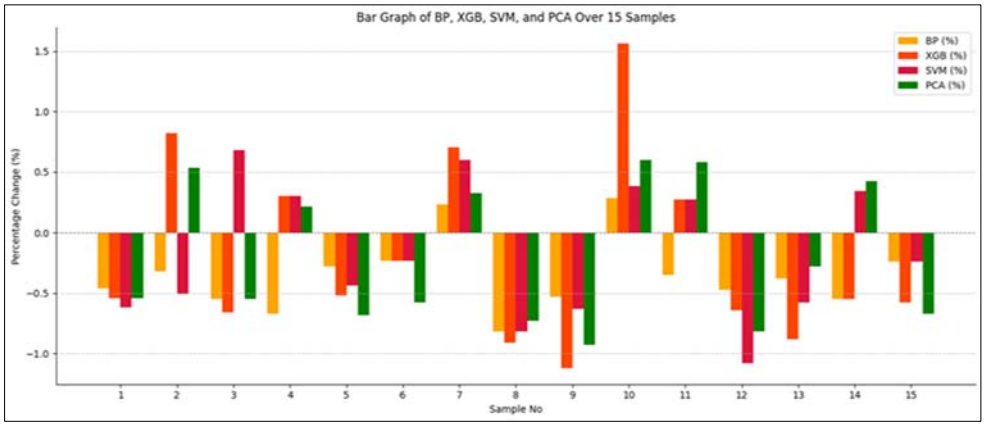

Fig. 10. Comparison plot on hardness

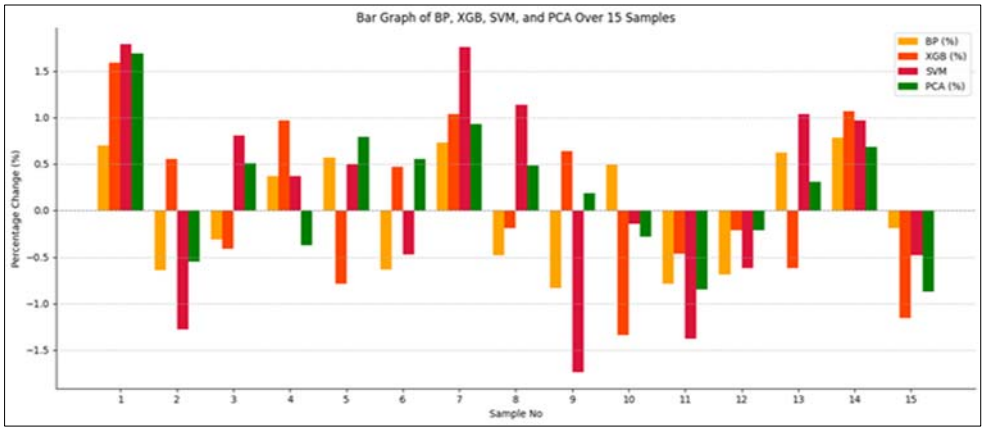

Fig. 11. Comparison plot on wear rate

Among all models and samples, the XGBoost model in Sample 12 showed the highest positive percentage change, exceeding +2.5%, indicating the highest overestimation. Conversely, the PCA model in Sample 13 registered the most negative error (–1.5%). BP and XGB exhibited more frequent and higher-amplitude variations across the Samples, especially samples 2, 9, 12, and 15. In contrast, SVM and PCA maintain relatively lower fluctuations in most samples, although PCA has a sharp negative spike in Sample 13. Overall, SVM is the best-performing model, which provided the lowest and most consistent prediction error across the samples for % of Elongation.

Figure 10 illustrates the percentage change in prediction error for four machine learning models across 15 samples for hardness. Among the four models, XGBoost in Sample 10 exhibited the highest positive error, reaching around +1.6%, and SVM in Sample 12 showed the highest negative error, around 1.1%. BP and XGB models display relatively higher variability, with several samples showing error magnitudes exceeding ±0.8%, especially in samples 2, 8, 9, 10, and 12, indicating less consistent predictions.

The PCA model maintained relatively low and balanced error values across most samples, suggesting greater stability and lower average prediction error. From these results, PCA was the most stable and accurate model in hardness, followed by SVM, BP and XGB.

Figure 11 displayed the percentage change in prediction errors of four machine learning models across 15 samples for wear rate. Among all the data points, the highest percentage error was observed in SVM for Sample 7, reaching approximately +1.7% and the lowest (most negative) error occurs in SVM again for Sample 9, around –1.6%. BP and XGB show moderate fluctuations across samples but with occasional peaks, such as BP in Sample 4 (+0.9%) and XGB in Sample 15 (–1.2%), highlighting variability. PCA maintained moderate and consistent performance, with errors mostly within ±1%.

4.3. Performance Metrics Results

The overall Performance metrics results for UTS, % of elongation, The hardness and wear rate results are displayed in Tables 1 to 4 respectively. Figure 12 presents a comparative analysis of R^2 (coefficient of determination) values for four machine learning models on output responses. Among all models and parameters, the highest R^2 value was observed for UTS using the SVM model, approximately 0.999, indicating nearly perfect prediction. Conversely, the lowest R^2 value was identified for Elongation (%) using the XGB

model, at around 0.989. Overall, BP and SVM models show strong and consistent performance across all four parameters, with R^2 values exceeding 0.996 in most cases and PCA performed well for wear rate and hardness.

Figure 13 illustrates the comparison of MAE for four machine learning models with four output responses. MAE represented the difference between predicted and actual values, where lower values indicate better model accuracy. Among all combinations, the highest MAE was detected for XGB in UTS, with an error close to 1.0, indicating the least accurate performance for this parameter.

Conversely, the lowest MAE was found in BP for Elongation (%), at approximately 0.09, showing high precision in predicting ductility. SVM shows the best performance with the lowest MAE (~0.65) for UTS. In terms of Hardness prediction, BP yields the lowest MAE (~0.29), whereas PCA showed the maximum error around 0.49. Notably, all models show low error in predicting Elongation %, but the PCA and XGB models consistently showed higher errors across multiple parameters in MAE. MSE is the squared differences between predicted and actual values.

From figure 14, the highest MSE was observed for XGB predicting UTS, with a value of 1.17, suggesting this model performs poorly in prediction of UTS. In contrast, the lowest MSE was attained for BP on Elongation (%), approximately 0.01, highlighting its exceptional accuracy in predicting ductility. For UTS prediction, SVM achieves the lowest MSE (~0.48), beating other models. When predicting Hardness, BP shows superior performance (~0.27 MSE) and PCA exhibits the highest error (~0.46). SVM performs worst with the highest MSE (~0.96), whereas BP shows relatively lower error (~0.58) for Wear Rate prediction.

Figure 14 shows RMSE prediction for four models. XGB showed the maximum RMSE for UTS (~1.07). Conversely, Wear Rate prediction yielded the low RMSE across all models (~0.005), suggesting minimal error and high predictive precision. For Elongation (%), BP performs best with the lowest RMSE (~0.09) and XGB exhibited the highest error (~0.18). For hardness, BP achieved the lowest RMSE (~0.13); however, PCA showed ~0.31. The consistently low RMSE values for these models indicate effective learning and stronger generalization capability, making them scientifically robust choices for predictive modelling.

Figure 15 shows MAPE prediction for four models and SVM achieves the lowest MAPE (~0.24%). BP showed the best performance (~0.76%), whereas XGB again yielded the highest error (~1.36%), indicating difficulty in modelling this response. For Hardness, BP outperformed other models with the lowest MAPE (~0.36%), whereas PCA shows the highest error (~0.55%). Wear Rate predictions maintain exceptionally low MAPE values across all models (~0.005–0.01%), confirming their consistent and highly accurate performance in this domain. Overall, BP emerges as the most balanced and accurate model, demonstrating the lowest average MAPE across multiple parameters, while XGB consistently records the highest errors, especially for Elongation and Hardness.

Table 1. Performance metrics results for UTS

Metric	BP	XGB	SVM	PCA
R^2	0.9981	0.9966	0.9986	0.9972
MAE	0.7420	0.9967	0.6453	0.9180
MSE	0.6280	1.1681	0.4838	0.9590
RMSE	0.7924	1.0808	0.6956	0.9793
MAPE	0.27	0.36	0.23	0.33
Accuracy	99.73%	99.64%	99.77%	99.67%

Table 2. Performance metrics results for % of elongation

Metric	BP	XGB	SVM	PCA
R^2	0.9970	0.9893	0.9937	0.9937
MAE	0.0887	0.1553	0.1247	0.1273
MSE	0.0085	0.0300	0.0176	0.0178
RMSE	0.0921	0.1733	0.1325	0.1333
MAPE	0.76	1.37	1.06	1.13
Accuracy	99.24%	98.63%	98.94%	98.87%

Table 3. Performance metrics results for hardness

Metric	BP	XGB	SVM	PCA
R^2	0.9971	0.9967	0.9967	0.9961
MAE	0.2873	0.4213	0.4267	0.4873
MSE	0.2608	0.3842	0.3982	0.4551
RMSE	0.1323	0.2383	0.2429	0.3088
MAPE	0.3638	0.4881	0.4928	0.5558
Accuracy	99.72%	99.68%	99.68%	99.59%

Table 4. Performance metrics results for wear rate

Metric	BP	XGB	SVM	PCA
R^2	0.998552	0.997312	0.996007	0.998238
MAE	0.000069	0.000089	0.000109	0.000071
MSE	0.587575	0.766869	0.967846	0.617114
RMSE	0.00023	0.00041	0.00058	0.00069
MAPE	0.000074	0.0001	0.000122	0.000081
Accuracy	99.43%	99.21%	99.03%	99.38%

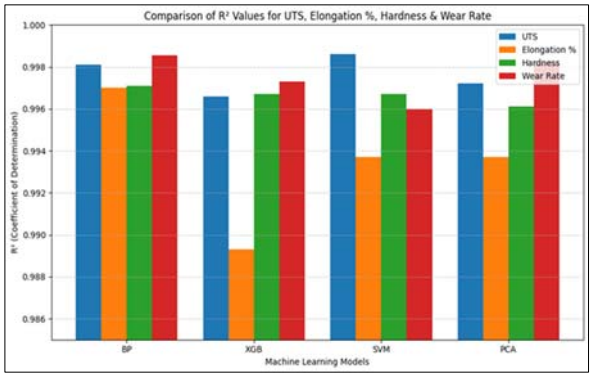

Fig. 12. Comparison plot of the coefficient of determination

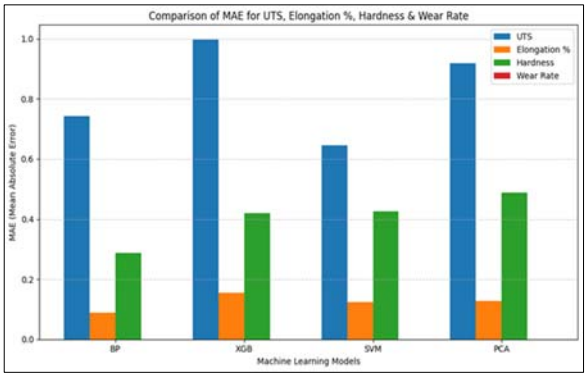

Fig. 13. Comparison plot of MAE

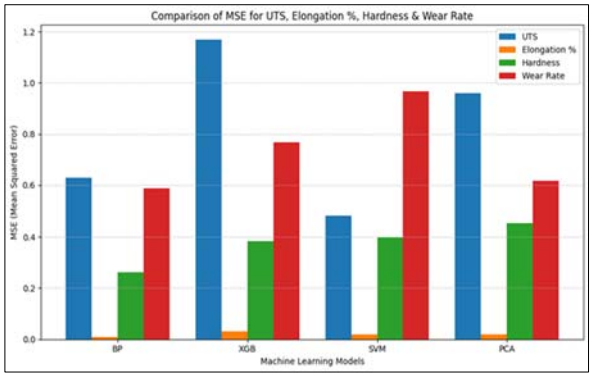

Fig. 14. Comparison plot of MSE

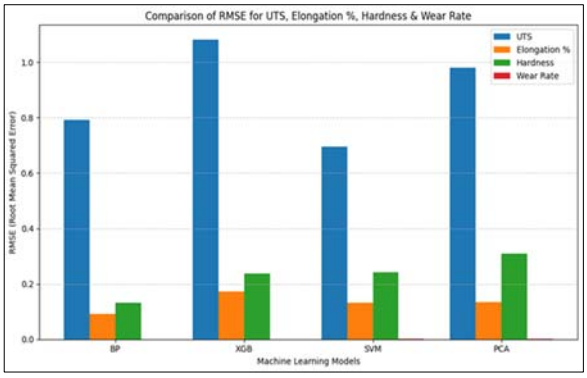

Fig. 15. Comparison plot of RMSE

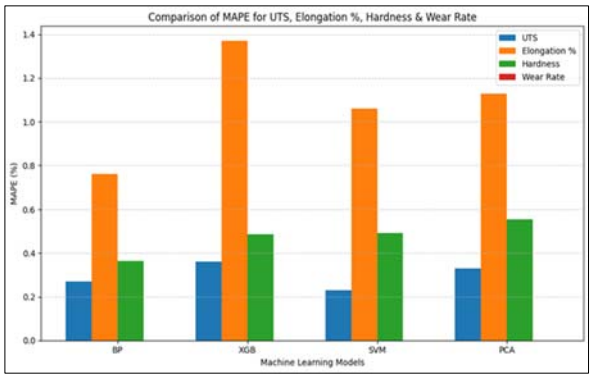

Fig. 16. Comparison plot of MAPE

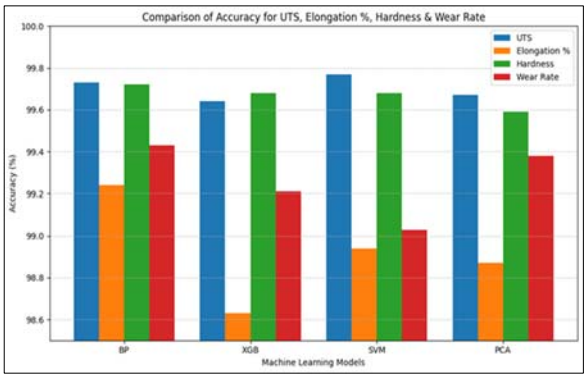

Fig. 17. Comparison plot of the accuracy percentage

Figure 15 shows MAPE prediction for four models and SVM achieves the lowest MAPE (~0.24%). BP showed the best performance (~0.76%), whereas XGB again yielded the highest error (~1.36%), indicating difficulty in modelling this response. For Hardness, BP outperformed other models with the lowest MAPE (~0.36%), whereas PCA shows the highest error (~0.55%). Wear Rate predictions maintain exceptionally low MAPE values across all models (~0.005–0.01%), confirming their consistent and highly accurate performance in this domain. Overall, BP emerges as the most balanced and accurate model, demonstrating the lowest average MAPE across multiple parameters, while XGB consistently records the highest errors, especially for Elongation and Hardness.

Figure 16 presents a comparative analysis of prediction accuracy (%) across four machine learning model for estimating four responses higher accuracy implied best model performance in approximating real-world outcomes. Among the models, SVM achieved the highest UTS prediction accuracy (~99.76%), followed by BP (~99.72%), XGB (~99.65%).

From figure 17, it can noticed that for elongation (%), BP leads again (~99.25%), whereas XGB showed the lowest accuracy (~98.63). In hardness prediction,

BP and XGB exhibited similar top-tier accuracy (~99.72%), but PCA yielded the lowest (~99.58%). Concerning with Wear Rate, BP stands out with the highest accuracy (~99.44%). These observations underlined that BP constantly delivered superior accuracy across all parameters, reflecting its robust learning capability and adaptability to nonlinear patterns in materials data.

5. CONCLUSIONS

This research explored the predictive modelling of Friction Stir Welding (FSW) output parameters for aluminium alloys using four advanced machine learning techniques: Backpropagation Neural Network (BP), Extreme Gradient Boosting (XGB), Support Vector Machine (SVM), and Principal Component Analysis (PCA)-enhanced regression. A carefully selected experimental dataset consisting of 15 observations was utilized to train and evaluate these models, targeting welding output responses such as Ultimate Tensile Strength (UTS), Elongation (%), Hardness, and Wear Rate.

- The correlation analysis revealed that the D/d ratio, tilt angle, and weld speed had the most significant influence on welding outcomes, whereas the input features demonstrated minimal intercorrelation for applying multivariate models like PCA.
- Actual versus predicted plots illustrated that all four models demonstrated strong predictive capabilities, with BP and XGB closely aligning with actual values for UTS, while SVM and PCA performed better in hardness prediction. For wear rate, BP and PCA exhibited more consistent alignment, especially at extreme values.
- In terms of error analysis, SVM consistently produced the lowest and most stable error percentages in predicting UTS and elongation, indicating its superior generalization and robustness. PCA emerged as the most reliable model for hardness prediction, maintaining balanced and low error margins across all samples. Although XGB and BP showed strong predictive power, they also exhibited higher variability, particularly in samples with extreme values, suggesting sensitivity to local fluctuations in the data.
- From the comparative accuracy results, SVM achieved the highest UTS accuracy (~99.76%), while BP led in elongation (~99.25%) and wear rate (~99.44%). PCA, although slightly lower in accuracy for some metrics, offered the most stable predictions, especially for hardness. Overall, the study confirms that ML-based models-particularly SVM and BP can effectively capture the complex nonlinear relationships inherent in the FSW process. PCA contributes significantly by enhancing model stability and reducing dimensionality issues. These findings provide a strong foundation for integrating ML techniques into real-time welding process monitoring, control, and optimization in advanced manufacturing settings.

REFERENCES

[1] **Wang X., Pan Y., Lados D. A.,** *Friction Stir Welding of Dissimilar Al/Al and Al/Non-Al Alloys: A Review*, Metallurgical and Materials Transactions B, 2018, vol. 49, pp. 2097–2117.

[2] **Sivasundar V. et al.,** *Predicting wear performance of AL6063 hybrid composites reinforced with Multi-Ceramic particles using experimental and ANFIS approaches*, Engineering Reports, 2025, vol. 7.

[3] **Chadha U. et al.,** *A Survey of Machine Learning in Friction Stir Welding, including Unresolved Issues and Future Research Directions*, Materials and Design Processing Communications, 2022, vol. 2022, pp. 1–28.

[4] **Reynolds A. P., Lockwood W. D., Seidel T. U.,** *Processing-Property Correlation in Friction Stir Welds*, Materials Science Forum, 2000, vol. 331–337, pp. 1719–1724.

[5] **Aziz S. B., Dewan M. W., Huggett D. J., Wahab M. A., Okeil A. M., Warren Liao T.,** *Impact of Friction Stir Welding (FSW) Process Parameters on Thermal Modeling and Heat Generation of Aluminum Alloy Joints*, Acta Metallurgica Sinica (English Letters), 2016, vol. 29, pp. 869–883.

[6] **Ahmadi H., Mostafa Arab N. B., Ghasemi F. A.,** *Optimization of process parameters for friction stir lap welding of carbon fibre reinforced thermoplastic composites by Taguchi method*, Journal of Mechanical Science and Technology, 2014, vol. 28, pp. 279–284.

[7] **Vijayakumar S.,** *Optimization of friction stir welding parameters for dissimilar aluminium alloys using RSM-GRA and RSM-TOPSIS: Towards Sustainable Manufacturing in Industry 4.0*, Results in Engineering, 2025, 107054.

[8] **Rajendran C., Srinivasan K., Balasubramanian V., Balaji H., Selvaraj P.,** *Identifying the combination of friction stir welding parameters to attain maximum strength of AA2014-T6 aluminum alloy joints*, Advanced Materials and Processing Technologies, 2018, vol. 4, pp. 100–119.

[9] **Tamjidy M., Baharudin B., Paslar S., Matori K., Sulaiman S., Fadaeifard F.,** *Multi-Objective Optimization of Friction Stir Welding Process Parameters of AA6061-T6 and AA7075-T6 Using a Biogeography Based Optimization Algorithm*, Materials, 2017, vol. 10, 533.

[10] **Abbasi M., Bagheri B., Keivani R.,** *Thermal analysis of friction stir welding process and investigation into affective parameters using simulation*, Journal of Mechanical Science and Technology, 2015, vol. 29, pp. 861–866.

[11] **Prabhakar D. A. P., et al.,** *A Review of Optimization and Measurement Techniques of the Friction Stir Welding (FSW) Process*, Journal of Manufacturing and Materials Processing, 2023, vol. 7, 181.

[12] **Arif S. et al.,** *Design, Development, Testing of Machine Learning Models to Estimate Properties of Friction Stir Welded Joints*, Materials, 2024, vol. 18, 94.

[13] **Okuyucu H., Kurt A., Arcaklioglu E.,** *Artificial neural network application to the friction stir welding of aluminum plates*, Materials and Design, 2007, vol. 28, pp. 78–84.

[14] **Shehabeldeen T. A., Elaziz M. A., Elsheikh A. H., Zhou J.,** *Modeling of friction stir welding process using adaptive neuro-fuzzy inference system integrated with harris hawks optimizer*, Journal of Materials Research and Technology, 2019, vol. 8, pp. 5882–5892.

[15] **Sahu P. K., Pal S.,** *Multi-response optimization of process parameters in friction stir welded AM20 magnesium alloy by Taguchi grey relational analysis*, Journal of Magnesium and Alloys, 2015, vol. 3, pp. 36–46.

[16] **Rabe P., Reisgen U., Schiebahn A.,** *Non-destructive evaluation of the friction stir welding process, generalizing a deep neural defect detection network to identify internal weld defects across different aluminum alloys*, Welding in the World, 2023, vol. 67, pp. 549–560.

[17] **Kumaresan S., Aultrin K. S. J., Kumar S. S., Anand M. D.,** *Transfer Learning With CNN for Classification of Weld Defect*, IEEE

Access, 2021, vol. 9, pp. 95097–95108.
[18] **Stephen D., Lalu P. P.,** *Development of Radiographic Image Classification System for Weld Defect Identification using Deep Learning Technique*, International Journal of Scientific and Engineering Research, 2021, vol. 12, pp. 390–394.
[19] **Fuse K., Venkata P., Reddy R. M., Bandhu D.,** *Machine learning classification approach for predicting tensile strength in aluminium alloy during friction stir welding*, International Journal of Interactive Design and Manufacturing, 2025, vol. 19, pp. 639–643.
[20] **Myśliwiec P., Kubit A., Szawara P.,** *Optimization of 2024-T3 Aluminum Alloy Friction Stir Welding Using Random Forest, XGBoost, MLP Machine Learning Techniques*, Materials, 2024, vol. 17, 1452.
[21] **Elsheikh A. H.,** *Applications of machine learning in friction stir welding: Prediction of joint properties, real-time control and tool failure diagnosis*, Engineering Applications of Artificial Intelligence, 2023, vol. 121, 105961.
[22] **Sarsilmaz F., Kavuran G.,** *Prediction of the optimal FSW process parameters for joints using machine learning techniques*, Materials Testing, 2021, vol. 63, pp. 1104–1111.
[23] **Rao T. V., et al.,** *Multi-response optimization of FSW parameters for Al-Mg-Zn alloys using Box-Behnken design and gray relational analysis and comparative study with ANFIS technique*, AIP Advances, 2025, vol. 15.
[24] **Liu H., Yu Q., Wu Q.,** *PID Control Model Based on Back Propagation Neural Network Optimized by Adversarial Learning-Based Grey Wolf Optimization*, Applied Sciences, 2023, vol. 13, 4767.
[25] **Chen Y., Zhang Y., Li C., Zhou J.,** *Application of XGBoost Model Optimized by Multi-Algorithm Ensemble in Predicting FRP-Concrete Interfacial Bond Strength*, Materials, 2025, vol. 18, 2868.
[26] **Abdi H., Williams L. J.,** *Principal component analysis*, WIREs Computational Statistics, 2010, vol. 2, pp. 433–459.
[27] **Waqar M., Dawood H., Guo P., Shahnawaz M. B., Ghazanfar M. A.,** *Prediction of Stock Market by Principal Component Analysis*, IEEE CIS Conference, 2017, pp. 599–602.
[28] **Wu Y., Zhou Y.,** *Hybrid machine learning model and Shapley additive explanations for compressive strength of sustainable concrete*, Construction and Building Materials, 2022, vol. 330, 127298.
[29] **Hu J., Bi J., Liu H., Li Y., Ao S., Luo Z.,** *Prediction of resistance spot welding quality based on BPNN optimized by improved Sparrow search algorithm*, Materials, 2022, vol. 15, 7323
[30] **Wu Z., Zhang Z., Song G.,** *Laser-ARC welding adaptive model of Multi-Pre-Welding condition based on GA-BP neural network*, Metals, 2025, vol. 15, p. 611.
[31] **Yu F., Zhao Y., Lin Z., Miao Y., Zhao F., Xie Y.,** *Prediction of mechanical properties and optimization of friction stir welded 2195 aluminum alloy based on BP neural network*, Metals, 2023, vol. 13, p. 267.
[32] **Wang R., et al.,** *CSSA-XGBoost: a novel algorithm for inspecting spot welding quality based on dynamic resistance signal*, Journal of Intelligent Manufacturing, 2025.
[33] **Vinodh D., Lakshmaiya N.,** *Integration of ceramic reinforcements in AA5083 composites for enhanced mechanical and thermal properties in friction stir welding*, Engineering Research Express, 2025, vol. 7, 035519.