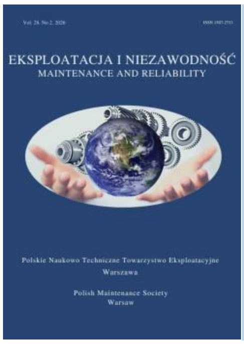


Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 28 (2026), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Li L, Jiao Y, Li X, Zhang Y, Liu L, Chen G, Song D, Prediction of torsional characteristics of clutch driven disc assemblies based on machine learning methods, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2026: 28(2) http:/10.17531/ein/215705

Prediction of torsional characteristics of clutch driven disc assemblies based on machine learning methods

Indexed by: Web of Science Group

Lijia Li[a], Yuan Jiao[a], Xiangyu Li[a], Yong Zhang[b], Liang Liu[c], Guihu Chen[b], Dunlan Song[a,*]

[a] Changchun University of Technology, China
[b] Zhejiang Qidie Automotive Parts Co., Ltd, China
[c] Jilin Zhonglian Testing Machine Manufacturing Co., Ltd, China

Highlights

- Predicting the torsional characteristics of clutch driven disc.
- Incremental learning opti-mization.
- Artificial Neural Networks.

Abstract

A method for predicting the torsional characteristics of clutch driven disc assemblies using machine learning algorithms is proposed. Traditionally, torsional stiffness is measured with a dedicated torsional testing machine. In this study, we aim to predict the torsional stiffness of various driven disc models based on historical testing data. Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN) models were employed for prediction. The results show that the ANN achieved the highest prediction accuracy, with R2=0.9590 and RMSE = 5.471 N·m/°. Furthermore, after optimization through incremental learning, the performance of the ANN model was significantly improved, achieving R2=0.9891 and RMSE = 3.045 N·m/°. This study substantially reduces the time and cost associated with measuring the torsional stiffness of clutch driven discs and demonstrates the potential of machine learning approaches in traditional mechanical engineering applications.

Keywords

artificial neural networks, incremental learning, random forest, torsional stiffness, clutch driven disc.

1. Introduction

The clutch driven disc serves as a vital transmission component, interfacing the automobile engine and gearbox. Its primary functions encompass transmitting engine torque and disengaging power during vehicle startup and gear shifts [1]. The torsional stiffness of the driven disc assembly has a direct impact on the clutch's transmission performance. Increased torsional stiffness reduces the deformation of the driven disc under torsional forces, thereby enhancing torque transmission efficiency [2,3].

This is essential for reliable clutch operation under various conditions, especially at high torque or speeds, where a stable transmission system ensures more dependable performance. Moreover, increased torsional stiffness minimizes driven disc deformation during torsion, thereby enhancing the clutch's response speed [4]. During vehicle startup, gear shifting, or acceleration, the clutch must swiftly transmit torque. Higher torsional stiffness minimizes transmission system delays, ensuring a quicker clutch response. Furthermore, the torsional

(*) Corresponding author.
E-mail addresses: L. Li (ORCID: 0009-0007-1507-5522) lilijia@ccut.edu.cn, Y. Jiao (ORCID: 0009-0005-6789-3557) 15836232352@163.com, X. Li (ORCID: 0009-0004-4065-7450) lxy1029427211@outlook.com, Y. Zhang (ORCID: 0009-0003-1731-5551) 13758158825@163.com, L. Liu (ORCID: 0009-0000-4469-4268) zhonglian_liul@126.com, G. Chen (ORCID: 0009-0003-8841-7934) 15857172657@163.com, D. Song (ORCID: 0009-0008-3395-224X) songdunlan@ccut.edu.cn

stiffness of the driven disc assembly plays a key role in reducing vibration and noise within the clutch system [5]. Appropriate torsional stiffness ensures the proper functioning of the transmission system, reducing accident likelihood [6]. In practical production, the factors influencing the torsional stiffness of the driven disc are multifaceted. Various parameters of the driven disc, such as material selection, geometric structure, and manufacturing precision, directly impact its torsional stiffness [7,8]. Figure 1 shows a clutch driven disc assembly with a diameter of 430 mm, including parts such as the friction plate, torsional damper, driven disc hub, and damping springs.

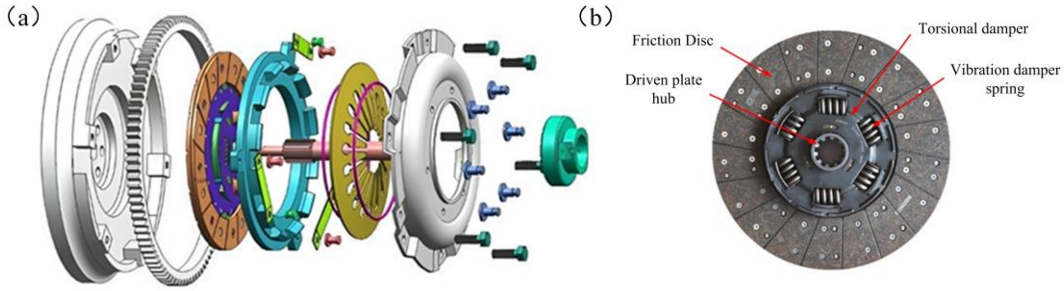


Fig 1. (a) Clutch Assembly; (b) Driven Plate Assembly.

In the early 20th century, measuring the torsional stiffness of driven discs relied primarily on mechanical devices and manual measurements. These devices utilized the lever principle and a ruler to gauge the angular deformation of materials under torsional force. Despite their simplicity, these methods laid the groundwork for future technological advancements. With the advent of electronic technology in the mid-20th century, electronic sensors and data acquisition systems were introduced, enhancing the precision and automation of torsional stiffness measurements. Strain gauges and rotary encoders became widely used in testing, significantly improving data accuracy and repeatability. [9] Recently, advancements in artificial intelligence and big data technologies have revolutionized the testing and analysis of torsional stiffness. Machine learning and data mining techniques now extract valuable insights from extensive test data, predict component behavior, and optimize design and manufacturing processes. [10-12]

Georgiou et al. [13] leveraged the advanced nonlinear capabilities of Artificial Neural Networks (ANN) to predict the response characteristics of flush end-plate beam-column connections (FEPC). They trained the ANN using approximately 200 samples from previous experimental programs, aiming to predict the bilinear response of FEPC, including strain hardening. This study described the ANN architecture, algorithms, and accuracy metrics of the new model. To determine the optimal design parameters for FEPC, Genetic Algorithms (GA) and Particle Swarm Optimization (PSO), two of the most accurate and widely used evolutionary optimization algorithms, were employed. In a separate study, Hu et al. [14] proposed a method to compensate for brake disc balancing errors using machine learning algorithms. By applying Random Forest, Decision Tree, and Support Vector Machine models to predict errors during the calibration process of balancing machines, the study explored methods to improve the balancing accuracy of brake discs in production. The results demonstrated that the Random Forest model achieved higher prediction accuracy compared to the Decision Tree and Support Vector Machine models.

Zeng et al. [15] proposed a neural network modeling technique based on Extreme Learning Machine (ELM) to predict automotive engine torque. This technique employs a single hidden-layer feedforward neural network (SLFN) structure, demonstrating the potential for high-accuracy approximation of any continuous function. Experimental results indicated that the model could accurately predict the torque of the gasoline engine. In practical applications, the proposed modeling method is expected to reduce engine testing and validation time, thereby significantly enhancing production efficiency. Jang et al. [16] proposed a method using deep learning algorithms to predict the strain history and durability assessment of automotive components. Their research employed Long Short-Term Memory (LSTM) networks to determine output data such as strain gauges and spring damping forces based on input data from wheel force and acceleration sensors. In recent years, researchers such as Özüpak et al. [17,18,20] have widely applied machine learning methods to various domains. They proposed several hybrid deep learning models (e.g., RNN–LSTM, MobileNetV2–ViT) and applied them to energy forecasting, air pollution prediction, and agricultural disease classification, while significantly improving model performance through techniques such as Bayesian optimization and cross-validation. Aslan et al. [19] integrated IoT with machine learning (e.g., LightGBM + grid search), enhancing both the accuracy and efficiency of power system prediction, and further incorporated SHAP and LIME to improve interpretability. Ansari et al. [21] focused on air quality prediction, employing multiple machine learning models for real-time AQI forecasting and identifying XGBoost as the most accurate and computationally efficient approach.

Based on the aforementioned results, this study further explores the potential of using machine learning models to predict the torsional stiffness of clutch driven discs. We collected extensive testing data for driven discs, including material properties, geometric dimensions, and torsional response data under different conditions. By performing detailed feature engineering on these data, we constructed various machine learning models, including Support Vector Machine (SVM), Random Forest, and Artificial Neural Network (ANN), to predict the torsional stiffness of driven discs. Furthermore, based on the continuous data obtained from subsequent testing by enterprises, we employed incremental learning methods to continuously optimize the existing machine learning models, thereby achieving higher prediction accuracy. The prediction process of the torsional stiffness of the clutch driven disc assembly based on machine learning methods is illustrated in Figure 2.

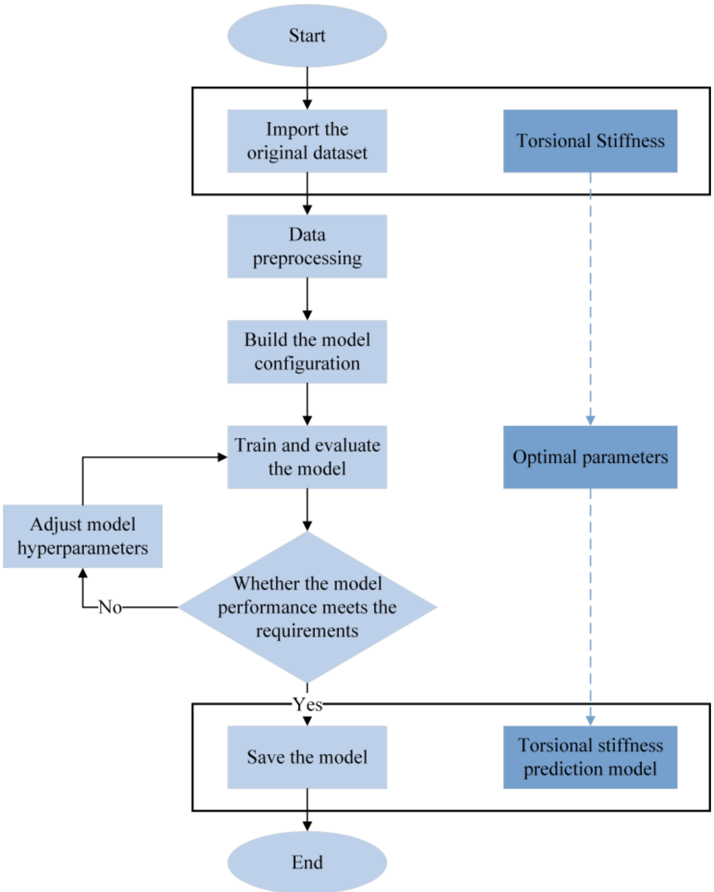


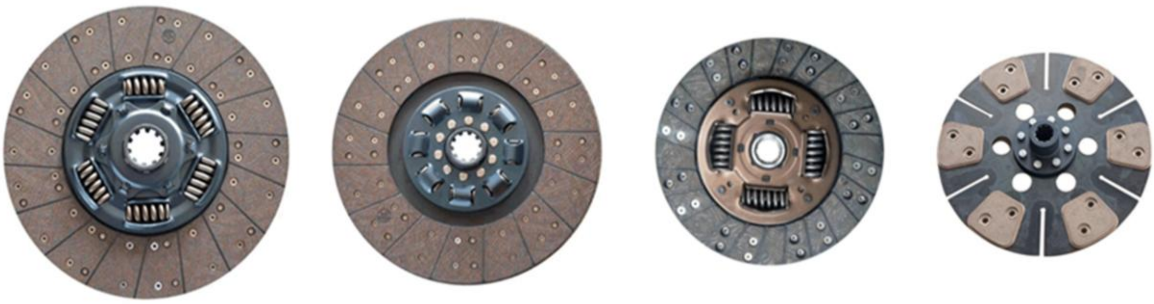
Fig 2. Torsional stiffness prediction process for clutch driven disc assemblies based on machine learning methods.

This study offers a scientific foundation for the design and optimization of clutch driven discs, while also showcasing the capabilities of machine learning, especially incremental learning, in traditional mechanical engineering domains. This interdisciplinary method enhances our understanding and prediction of variations in the torsional stiffness of driven discs, thus providing more efficient and reliable solutions for the automotive industry.

2. Research data

2.1. Acquisition and Processing of Experimental Data

The test data were derived from torsion characteristic tests of driven disc assemblies for four models (denoted as A, B, C, and D, respectively), totaling 138 groups. Schematic diagrams of the driven discs for the four models are shown in Figure 3 with their structures and related parameters presented in Table 1. Among the four models, Model A, corresponding to a driven disc assembly with a 430 mm diameter, has a higher volume of test data compared to the other three models. This discrepancy is primarily due to the widespread use of this model's driven disc in the current heavy-duty commercial vehicle sector. In the study of predicting the torsion performance of driven disc assemblies, conducting larger-scale data collection for this model is representative, offers significant engineering guidance, and facilitates the development of more universally applicable performance prediction models.



(a)A (b)B (c)C (d)D

Fig 3. Model of the Tested Driven Disc Assembly: （a）Model A, （b）Model B,（c）Model C, （d）Model D.

Table 1. Relevant Parameters of Driven Disc Assemblies of Different Models.

Driven Disc Model	Model of the Driven Disc	Diameter of the Driven Disc (mm)	Number of Damping Springs (units)	Number of Wave Plates (units)	Thickness of Wave Plate (mm)
A	430	205	6	9	0.80
B	395	184	8	8	0.70
C	300	150	4	6	0.65
D	280	68	0	6	0.55

Fig 4. Torsional Characteristic Test Rig for Clutch driven Disc Assembly.

The torsion characteristic test of the clutch driven disc assembly is primarily used to evaluate the response of the clutch driven disc assembly to torsional deformation during operation, ensuring that the performance of the clutch driven disc assembly in actual use meets the design requirements. This test simulates the torsional loads on the driven disc under various operating conditions, measuring parameters such as torsional stiffness, hysteresis characteristics, and the occurrence of torsional vibrations. The torsional stiffness testing equipment is a torsion characteristic testing machine from the automated inspection line for clutch driven disc assemblies, produced by a partner company of the project team, as shown in Figure 4. During the test, data are collected and recorded for subsequent analysis and evaluation, with products failing to meet the specified torsional stiffness range promptly marked as non-compliant.

The testing machine applies varying torque loads to simulate power transmission in a vehicle under different operating conditions. It uses torque sensors, angle sensors (or displacement sensors), and vibration sensors to monitor the torque response, angular displacement, and vibration of the clutch driven disc in real time, generating a torsion characteristic curve.

$$C_a = \frac{M_{max}}{\theta_{max}} \tag{1}$$

In this context, C_α is defined as the torsional stiffness, M_{max} is the ultimate torque capacity, and θ_{max} denotes the ultimate torsional angle. The torsion stiffness testing machine on the automated inspection line is used to measure the torsional stiffness of each product, storing the torsional stiffness values and other parameters of each driven disc in a data management system. These parameters include six items: the diameter of the driven disc assembly, the diameter of the torsion damper, the number of damping springs, the pre-damping angle, the preload torque, and the ultimate torque. The first three parameters are inherent properties of the driven disc. The preload torque is an initial torque applied to the driven disc by the testing machine before formal testing begins, which eliminates contact gaps between the sample and the equipment, stabilizes the contact surface, and reduces errors. The pre-damping angle refers to an initial torsional angle applied to the sample before formal testing, which eliminates initial slack and stabilizes material properties. The ultimate torque is the maximum torque applied by the testing machine during the test, varying across different models of driven discs.

The subsequent study will utilize the test data from the torsion stiffness testing machine on the automated inspection line for clutch driven disc assemblies, including applied test parameters such as torsion torque and pre-damping angle, as well as relevant parameters of the driven disc. The collected test data will undergo preprocessing steps, including data cleaning, data integration, and data transformation, to be consolidated into a training dataset for the predictive model.

2.2. Statistical Analysis of Data

Before modeling the machine learning model, a statistical analysis was conducted on the target variable (torsional stiffness of the driven disc assembly) to facilitate the evaluation of the model's prediction results. The statistical results for torsional stiffness are shown in Table 2. The test data, comprising 138 groups, were sourced from four different models of driven discs. To ensure the reliability of the test data, each driven disc assembly was subjected to multiple (no fewer than three) torsional stiffness tests, with the average of these tests recorded as the final data point for the sample. During the testing process, influencing factors such as ambient temperature, loading rate, and fixture clamping method were strictly controlled to minimize testing errors. The maximum torsional stiffness value of the driven disc assemblies was 497.17, observed in Model A, while the minimum was 42.69, observed in Model C. The mean torsional stiffness was 384.85, with a standard deviation of 170.79.

Table 2. The Torsional Stiffness Distribution of the Tested Clutch driven Disc Assemblies.

Target Variable	**Maximum Value** $(\frac{N\cdot m}{o})$	**Minimum Value** $(\frac{N\cdot m}{o})$	**Mean Value** $(\frac{N\cdot m}{o})$	**Median Value** $(\frac{N\cdot m}{o})$	**Standard Deviation** $(\frac{N\cdot m}{o})$
Torsional Stiffness	497.17	42.69	384.85	480.47	170.79

In designing the torsional stiffness prediction model for the clutch driven disc assembly, this study considered six key parameter variables: driven disc diameter, torsion damper diameter, number of damping springs, pre-damping angle, preload torque, and ultimate torque. Figure 5 illustrates the data distribution of each input feature variable for the four models of driven disc assemblies.

Due to manufacturing and assembly tolerances of the driven plate, the torsional stiffness obtained from tests on driven plates of the same model is not a fixed value but falls within the designed torsional stiffness range for each model. For Model A clutch driven plates, the torsional stiffness is concentrated in the range of 470 to 500. For Model B, it ranges from 420 to 460. Model C driven plate assemblies exhibit torsional stiffness values between 42 and 75, while Model D ranges from 80 to 105. Regarding the applied parameters in testing, even for driven plates of the same model, the applied maximum torque may vary due to differences in test objectives, equipment settings, and sample conditions. The maximum torque also differs significantly across models due to differences in design purpose and structural parameters. For Model A, the maximum torque during testing is approximately 4400; for Model B, around 3400; for Model C, approximately 950; and for Model D, about 650. Preload torque is used to eliminate gaps between components and ensure stability in the initial state. The pre-damping angle reflects the angular range at which the internal damping components of the driven plate begin to engage in torsional work, serving as a key indicator of the system's initial compliance. For Model A, the preload torque ranges from 64 to 80; for Model B, from 35 to 40; for Model C, from 7 to 10; and for Model D, from 65 to 80. The pre-damping angle is approximately 4°–4.5° for Model A, 3°–4.3° for Model B, 3°–4.5° for Model C, and 5.5°–7° for Model D. Based on the distribution of the data, it can be observed that the torsional stiffness follows certain patterns within specific input parameter ranges.

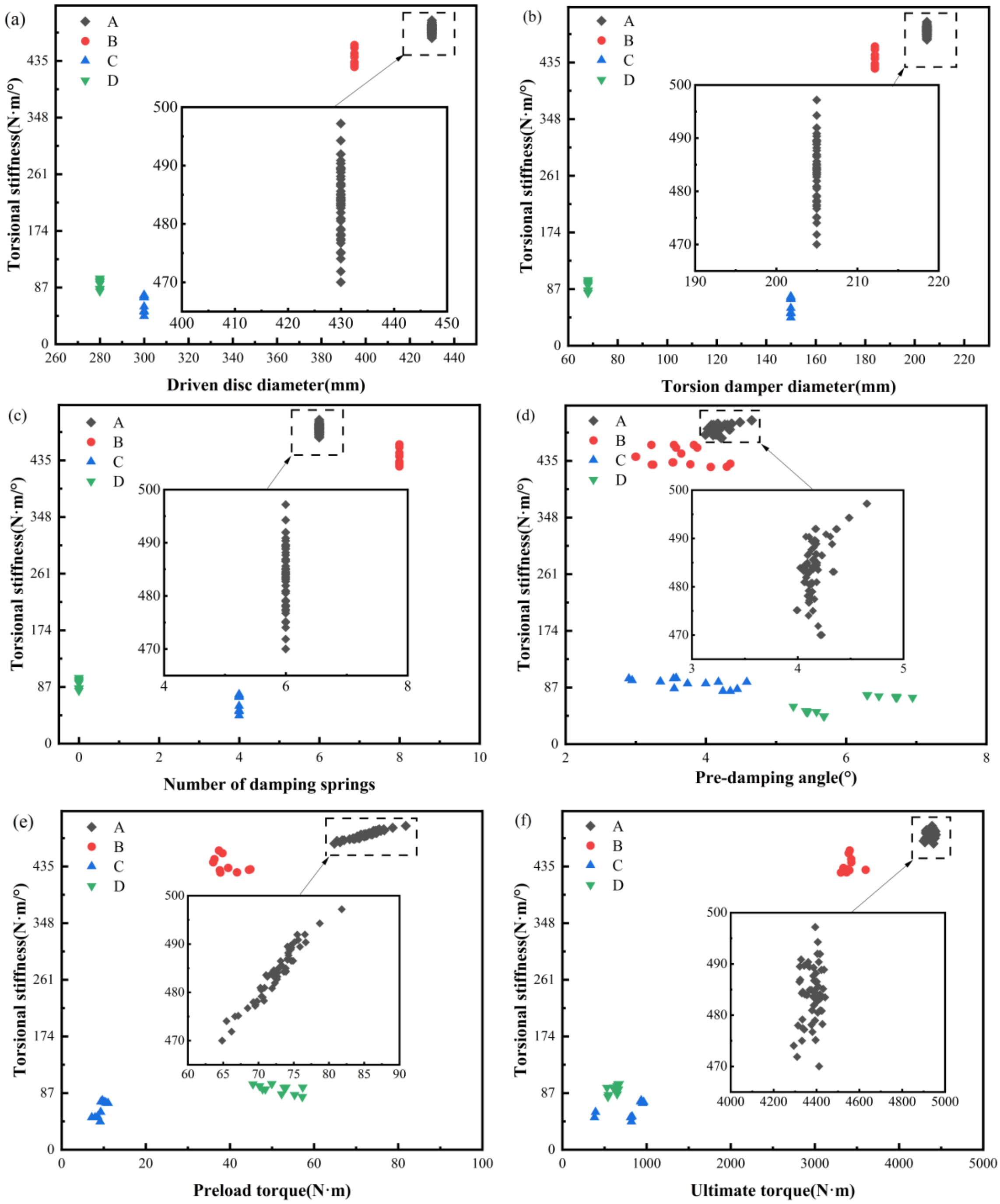


Fig 5. Data Distribution of Input Feature Variables for the Model: (a)Driven disc diameter, (b)Torsion damper diameter, (c)Number of damping springs, (d)Pre-damping angle, (e)Preload torque, (f)Ultimate torque.

To gain deeper insight into the relationships between input features and target outputs prior to machine learning modeling, this study employs Pearson correlation analysis to examine the inter-variable correlations within the dataset. A heatmap is used to visually present the pairwise correlations between variables. The Pearson correlation coefficient measures the degree of linear correlation between two continuous variables, ranging from -1 to 1. A value closer to 1 or -1 indicates a stronger correlation, while a value of 0 suggests no linear relationship. The formula for calculating the Pearson correlation coefficient is as follows:

$$r = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}} \tag{2}$$

In the formula, X_i and Y_i represent the values of the $i-th$ sample for features X and Y, respectively, while $\overline{X}$ and $\overline{Y}$ denote the mean values of the corresponding features. The correlation coefficient heatmap derived from the torsional stiffness test dataset of the driven plate assembly is shown in Figure 6.

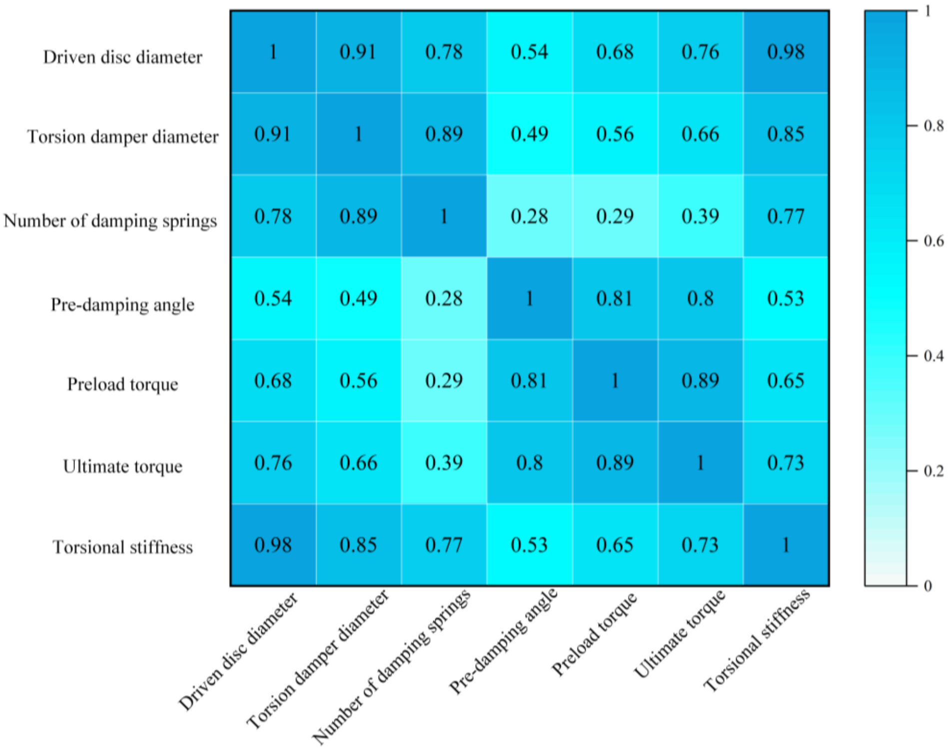


Fig 6. Pearson Correlation Coefficient Chart between Model Input Feature Variables and Torsional Stiffness.

The features in the dataset exhibit generally high correlations with the target variable (torsional stiffness), which is consistent with the practical reality of coordinated structural parameter design in clutch driven plate assemblies. Accordingly, the dataset provides a solid foundation for modeling and is well-suited for torsional stiffness prediction using regression-based machine learning models. However, the high degree of correlation among features may lead to multicollinearity issues in model interpretability. This is particularly problematic for traditional regression models such as linear regression, where the regression coefficients can become unstable. Machine learning methods that are less sensitive to feature correlation—such as Random Forest, Support Vector Machines (SVM), and Neural Networks—are better equipped to handle such scenarios. In the following sections, torsional stiffness prediction for driven plate assemblies will be conducted using these three machine learning models.

3. Machine learning methods

Machine learning is a data-driven algorithmic technique that enables computers to learn and recognize patterns from data without explicit programming instructions, thereby making predictions and decisions. Machine learning methods are generally categorized into three main types based on the characteristics of the data and the target tasks: supervised learning, unsupervised learning, and reinforcement learning. In this study, we primarily focus on the regression tasks within supervised learning, utilizing various machine learning algorithms to predict the torsional stiffness of clutch driven discs.

In machine learning, the goal is to learn a function f from data $X = \{x_1, x_2, \dots, x_n\}$ and corresponding labels or outputs $Y = \{y_1, y_2, \dots, y_n\}$ such that for a given input$x$, the output $y \approx f(x)$. [22]

To measure the discrepancy between the predicted values of the model f and the actual values, a loss function $L(y, \hat{y})$ is defined, where y is the actual value and $\hat{y} = f(x)$ is the predicted value. A commonly used loss function is the Mean Squared Error (*MSE*), defined as:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3}$$

where y_i is the i-th observed value, $\hat{y}_i$ is the $i = th$ predicted value by the model, and n is the total number of observations. The goal of machine learning is to find the optimal model parameters θ that minimize the loss function L:

$$\theta^* = argmin_\theta \, L\left(Y, f(X, \theta)\right) \tag{4}$$

where $f(X, \theta)$ represents the model's prediction for the input data X with parameters θ. During the model training and evaluation process, two commonly used metrics are and Root Mean Squared Error (*RMSE*). R^2 , the coefficient of determination, is used to evaluate the model's explanatory power. It is defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{5}$$

where $\bar{y}$ is the mean of the observed values. The value of R^2 ranges between 0 and 1, with values closer to 1 indicating stronger explanatory power of the model. The Root Mean Squared Error (*RMSE*) measures the standard deviation of the differences between the predicted values and the actual values. It is defined as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)} \tag{6}$$

A smaller *RMSE* value indicates higher prediction accuracy of the model. By using the aforementioned two evaluation metrics, R^2 and *RMSE*, one can comprehensively assess the performance of machine learning models in predicting the torsional stiffness of clutch driven discs. This provides a basis for selecting the optimal model. [23]

3.1. Support vector machines

Support Vector Machines (SVMs) can be applied to both classification and regression tasks, making them widely used in performance prediction. The fundamental concept of SVMs is to construct an optimal decision hyperplane that achieves effective classification. For linearly separable data, SVM enhances classification performance by selecting the hyperplane with the maximum margin [24]. For non-linearly separable data, kernel functions are employed to map the input into a higher-dimensional feature space, thereby enabling linear separability. A key advantage of SVMs lies in their strong classification capability in high-dimensional spaces. They demonstrate robustness and stability when handling complex and nonlinear datasets, as the decision boundary is determined only by the support vectors near the margin rather than the entire dataset [25]. Figure 7 illustrates the operating principle of SVMs.

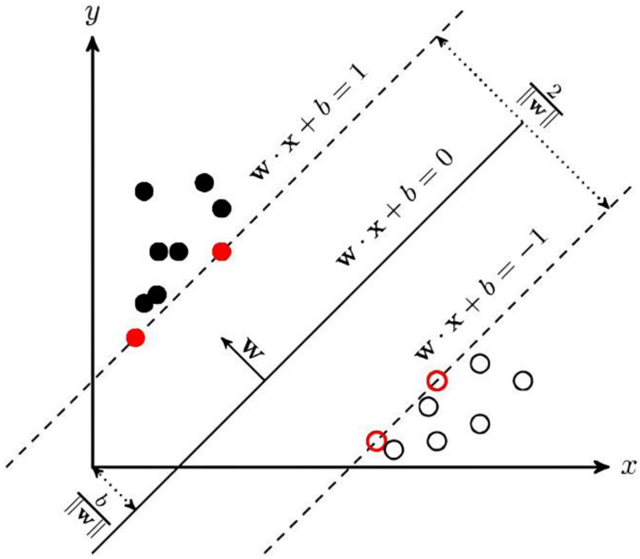


Fig 7. Schematic Diagram of the Support Vector Machine Principle.

3.2. Random forest

Random Forest (RF) is a widely used ensemble learning method applicable to classification, regression, and various other tasks. It enhances model accuracy and stability by constructing multiple decision trees and aggregating their predictions through majority voting (for classification) or averaging (for regression). The core principle of RF is to introduce randomness in the generation of independent decision trees and then combine their outputs, thereby reducing variance and mitigating the risk of overfitting [26]. The overall architecture of the Random Forest model is illustrated in Figure 8.

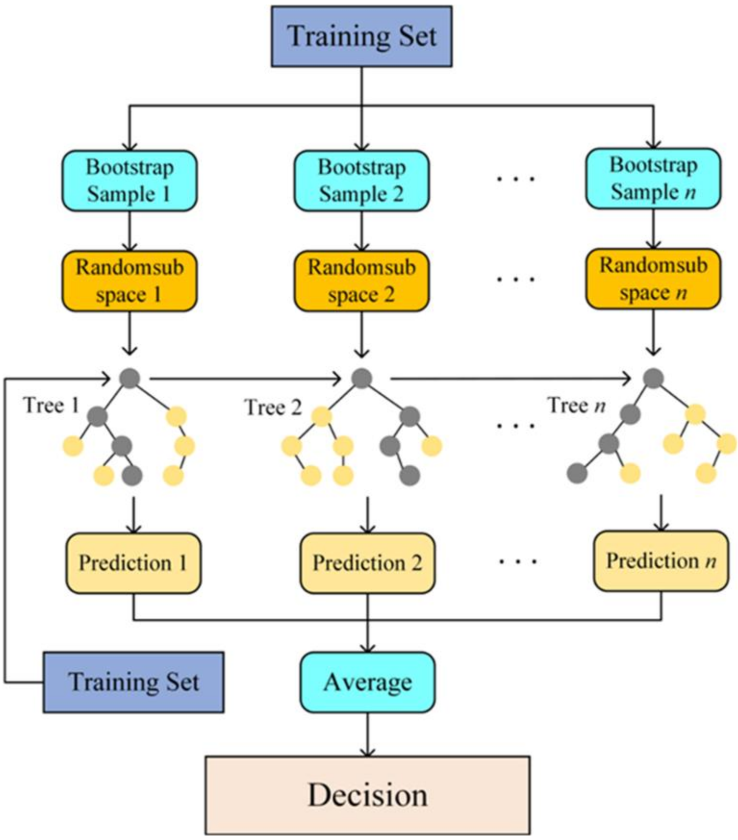


Fig 8. General Architecture Diagram of the Random Forest Model.

Several key parameters influence RF performance. First, the number of decision trees is critical: increasing the number of trees generally improves accuracy but also increases computational cost and memory usage. Second, the number of features randomly selected at each node split directly affects diversity and model generalization. In addition, restricting the maximum tree depth prevents overly complex trees and reduces the risk of overfitting. Other parameters, such as the minimum number of samples required to split a node or to remain in a leaf node, can also be tuned, where larger values help further mitigate overfitting. The bootstrap parameter determines whether sampling with replacement is used to generate training subsets and whether out-of-bag (OOB) samples are utilized to estimate generalization error. Finally, the number of CPU cores allocated for parallel training can be specified to improve computational efficiency [27].

3.3. Artificial Neural Network

Artificial Neural Networks (ANNs) are computational models inspired by the structure and function of biological neural networks. They are typically composed of an input layer, one or more hidden layers, and an output layer. Input features are processed through the hidden layers, where nonlinear relationships are captured using activation functions. With multiple hidden layers and a sufficiently large set of parameters, ANNs demonstrate strong adaptability and are capable of modeling complex mappings between inputs and outputs [28–30].

The primary hyperparameters of an ANN include the number of hidden layers and neurons per layer, the choice of activation function, the optimization algorithm, the learning rate, and the maximum number of training iterations. These hyperparameters directly influence convergence speed, predictive performance, and the stability of the training process. While ANNs possess powerful nonlinear modeling capabilities that enable them to adapt to complex data distributions, inappropriate hyperparameter settings may prolong training, increase computational cost, or trap the model in local optima, ultimately reducing predictive accuracy. Therefore, careful design of the network architecture and training strategy—particularly the configuration of hidden layers and the selection of learning parameters—is essential to improving ANN performance. A typical ANN structure with two hidden layers is illustrated in Figure 9.

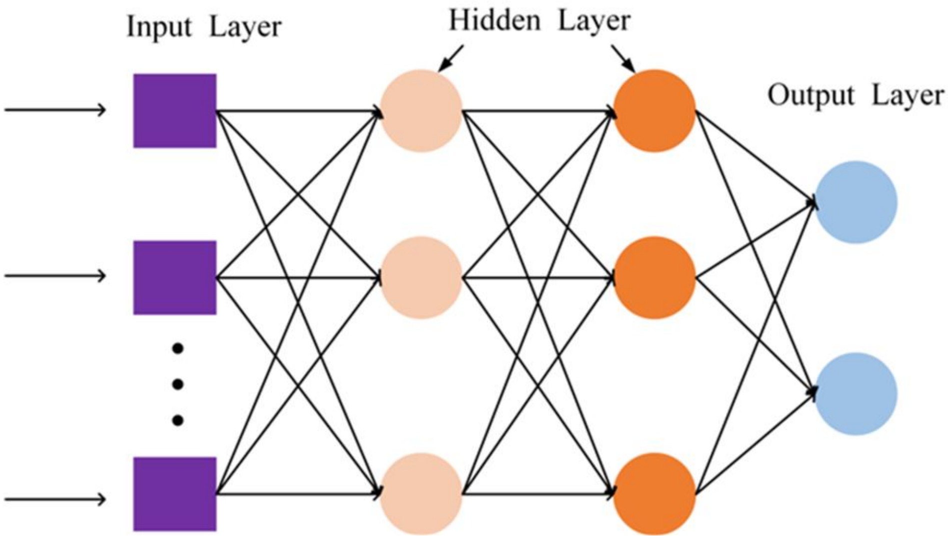


Fig 9. Diagram of the Structure of an Artificial Neural Network with Two Hidden Layers.

4. Model results and discussion

4.1. Model performance

Prior to fitting the three machine learning models, the corresponding datasets must undergo preprocessing. To ensure optimal predictive performance, the grid search method is employed to identify the best set of hyperparameters for each model. Grid search involves systematically exploring all possible combinations of hyperparameter values within a predefined range, evaluating each configuration to find the optimal one. During the search process, the model's performance is assessed using cross-validation, and the hyperparameter combination that yields the best results is selected.

To reduce computational burden and time costs, the hyperparameter search space for the Artificial Neural Network (ANN) model is appropriately constrained. For example, the initial learning rate is limited to values such as 0.001, 0.005, and 0.01, with similar bounded ranges applied to other parameters. Although this tuning approach requires substantial computation and may extend training time, it ensures high model performance by enabling a comprehensive and orderly exploration of all parameter combinations, ultimately identifying the most effective configuration for the given dataset.

The optimal parameters obtained from training the models are as follows: For the SVM model, the best parameters are: regularization parameter $C = 1$, kernel function parameter $\gamma = 1$, and the kernel type is the polynomial kernel. For the Random Forest (RF) model, the optimal settings include: maximum tree depth = 10, minimum number of samples at each leaf node = 2, minimum number of samples required to split an internal node = 2, and number of trees = 300. For the Artificial Neural Network (ANN) model, the optimal parameters are: the activation function for both hidden and output layers is the hyperbolic tangent function, the L2 regularization parameter is set to 0.01, the network consists of a single hidden layer with 100 neurons, the initial learning rate is 0.001, the maximum number of iterations is 1000, and the optimization algorithm used is stochastic gradient descent. The optimal parameter settings of the artificial neural network (ANN) model are shown in Table 3.

Table 3. Hyperparameter Settings of the Neural Network.

Hyperparameter Name		Preset Value
Number of Network Layers		3
Input Layer	Number of Neurons	6
	Activation Function	/
Hidden Layer	Number of Neurons	100
	Activation Function	Tanh Function
	Number of Neurons	1

Output Layer	Activation Function	Tanh Function
	Loss Function	*RMSE*
	Evaluation Metric	R^2
	Number of Iterations	1000
	Batch Size	3

The R^2 values, *RMSE* values, and training times of the three models on the same dataset are presented in Table 4. As shown, all three models exhibit relatively high R^2 values, indicating strong explanatory power and good fitting performance, with the ability to effectively utilize input features to predict the target variable. Among them, the ANN model achieves the best predictive performance, with the highest R^2 and the lowest *RMSE*, though it also requires the longest training time. The RF model ranks second, with prediction errors slightly higher than those of the ANN. The SVM model has the shortest training time, but its prediction error is the highest among the three.

Table 4. Performance Comparison of the Three Basic Models.

Model	R^2	***RMSE***	**CPU Time**
SVM	0.9133	10.323	270ms
RF	0.9388	6.483	2.6s
ANN	0.9590	5.471	5.4s

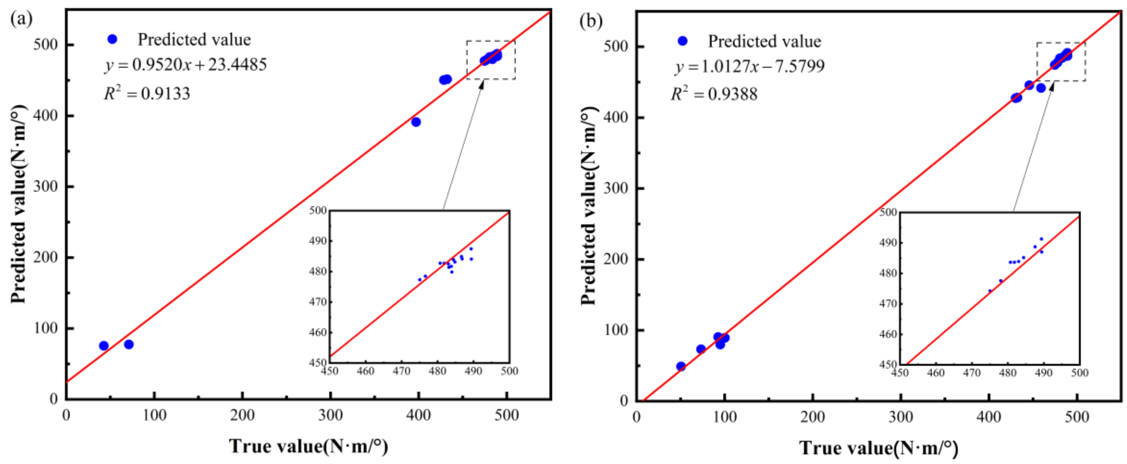
4.2. Comparison of prediction results

To verify the accuracy of the model predictions, the torsional stiffness values obtained from the torsion testing machine were used as the ground truth. The three models were used to predict the torsional stiffness of 20 randomly selected driven disc assemblies. The predicted torsional stiffness values from each model were then compared with the experimentally measured values. The prediction results of the three models are shown in Figure 10.

The SVM model achieved an R^2 of 0.9133, with a root mean square error (*RMSE*) of 10.323 N·m/°, and had the shortest training time among the three models. During prediction, the maximum error occurred when the predicted value exceeded the actual value by 32.426 N·m/°, while the minimum error was 0.57 N·m/°.

The RF model achieved an R^2 of 0.9388 and an *RMSE* of 6.483 N·m/°. Its maximum prediction error occurred when the predicted value was 17.776 N·m/° lower than the actual value, and the minimum error was 0.456 N·m/°.

The ANN model yielded the best performance, with an R^2 of 0.9590 and an *RMSE* of 5.471 N·m/°, though it also required the longest training time. The model's maximum prediction error was an underestimation of 16.475 N·m/°, and the minimum error was just 0.018 N·m/°.



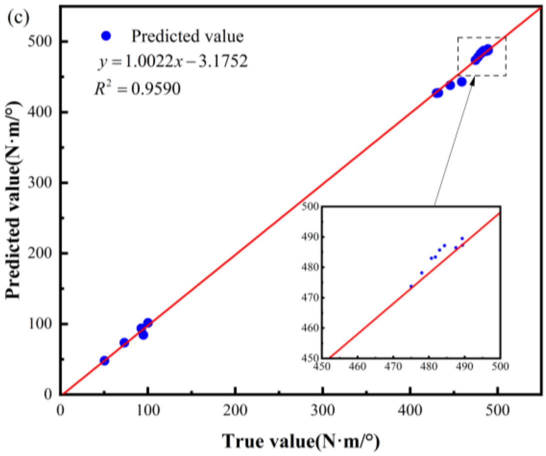


Fig 10. Torsional Stiffness Prediction Results Obtained from Three Different Basic Models: (a)SVM Model Prediction Results, (b)RF Model Prediction Results, (c)ANN Model Prediction Results.

Based on the comparison of prediction results from the three models, the ANN and RF models produce outputs that are closer to the actual values, while the performance of the SVM model is slightly inferior. This observation is consistent with the corresponding root mean square error (*RMSE*) values of the three models.

In conclusion, this study selects the Artificial Neural Network (ANN) model as the optimal model for predicting the torsional characteristics of clutch driven disc assemblies. Subsequent research will focus on incremental learning optimization for this model. With this predictive model, the torsional stiffness of four types of driven discs can be estimated based on their input parameters, thereby reducing the need for physical testing to some extent and saving both time and cost.

5. Incremental learning optimization

A common challenge in applying machine learning models to predict the mechanical performance of the driven disc is that, as experimental data continuously accumulates, the model must often be retrained to incorporate the new data, which usually leads to the abandonment of the previous model. This training strategy typically poses two major drawbacks: first, it demands substantial computational time and resources; and second, it incurs considerable costs [31]. To address these limitations, researchers have proposed the concept of incremental learning. Incremental learning enables a model to be locally updated and refined with the continuous influx of new data, without the need for complete retraining of the original model. Its core principle is to progressively adapt the model to new input data in an "incremental" manner, while effectively preserving the knowledge already acquired [32–34].

In neural network–based incremental learning, when new data are introduced, the model does not discard previously learned information. Instead, it assimilates the new data and continuously adjusts the network weights according to the evolving data distribution. This strategy allows the model to integrate new information while avoiding catastrophic forgetting of prior knowledge. Compared with conventional models trained on static datasets, incremental learning significantly improves computational efficiency. Consequently, it is particularly suitable for real-world scenarios where data evolves dynamically in real time and expands continuously, thereby providing strong practical applicability [35].

In the practical production and testing environment of clutch driven discs, data are often continuously changing, making it difficult to obtain a complete and ideal training dataset at once. Therefore, to ensure reliability in manufacturing applications, a predictive model must possess the capability for continuous learning and dynamically adapt to newly emerging data, thereby maintaining consistent predictive accuracy under varying conditions. Incremental learning provides an effective solution to this challenge. By performing localized updates to the model based on both the existing structure and the newly acquired training data, it eliminates the need for full retraining, substantially reducing computational resource consumption and shortening training time.

5.1. Model Optimization Process

Artificial Neural Network (ANN) models inherently support incremental learning, as their weights can be updated either in batches or on a sample-by-sample basis. This property makes ANNs highly flexible and particularly suitable for incremental learning applications.

In this chapter, the torsional characteristic prediction model developed in the preceding sections is further optimized using incremental learning. The new data for this optimization consist of 100 additional torsional characteristic tests conducted on clutch driven disc assemblies. A subset of the incremental dataset is presented in Table 5.

Table 5. Dataset Added by Incremental Learning (Partial).

No.	Driven Disc Diameter (mm)	Torsion Damper Diameter (mm)	Number of Damping Springs (units)	Pre-Damping Angle (°)	Preload Torque (N·m)	Ultimate Torque (N·m)	Torsional Stiffness (N·m/°)
1	430	205	6	4.15	67.16	4397.05	484.23
2	300	150	4	6.72	10.61	955.99	71.57
3	280	68	0	2.74	48.47	668.43	92.61
4	430	205	6	4.19	66.21	4310.67	475.12
5	395	184	8	2.83	37.41	3416.86	459.36

The dataset is divided into a training set and a validation set: 20% of the samples are reserved as the validation set to compare the predictive performance of the incrementally updated model with that of the original model, while the remaining 80% are used to construct the training–evaluation framework for the current incremental learning setup.

The incremental learning optimization method for artificial neural networks proposed in this chapter is primarily designed to address knowledge updates resulting from an increased volume of sample data. It does not consider learning related to the expansion of categories or attributes. The specific process of model optimization through incremental learning is illustrated in Figure 11 and consists of the following five steps:

(1) Original Model Loading: The system first loads the historically best-performing model architecture and weights as the starting point for incremental training.

(2) Incremental Data Preprocessing: Newly collected experimental data are normalized and outliers are removed to ensure the quality of the training data.

(3) Model Fine-Tuning: The cleaned new sample data are fed into the model for fine-tuning of specified network layers. In this stage, a strategy of freezing part of the network weights is applied, updating only the higher-level layers to avoid catastrophic forgetting.

(4) Performance Evaluation and Replacement Mechanism: After training, the updated model is compared with the original model in terms of R^2 and RMSE on the validation set. If the incremental model outperforms the original model on both metrics, the original model is replaced with the updated version; otherwise, the original model is retained.

(5) Model Saving and Version Control: All optimized models are saved as independent versions to facilitate backtracking and future model fusion.

5.2. Model Parameter Settings

After completing data cleansing and integrity checks on the incremental dataset, it is necessary to normalize the new batch of data using the statistical parameters obtained from the original training set. This procedure ensures consistency between the features of the new data and those used during the initial model training. Each feature is normalized according to the statistical metrics derived from the training dataset, thereby preserving the feature-wise distribution of the incremental data. Such consistency guarantees the comparability of each incremental batch and, as the volume of training samples increases, contributes to improving the predictive accuracy of the neural network model. To provide a visual demonstration of performance evolution during the incremental learning process, Figure 12 illustrates the trend of prediction error with respect to training epochs.

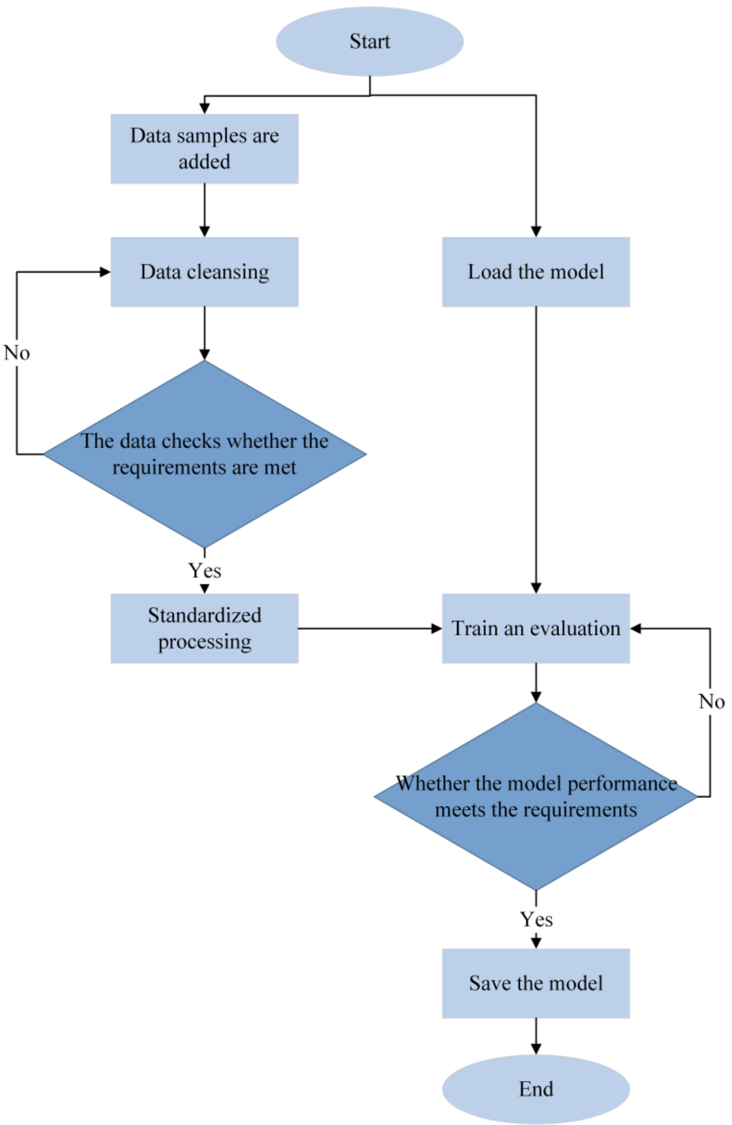


Fig 11. Flowchart of the Incremental Learning Algorithm.

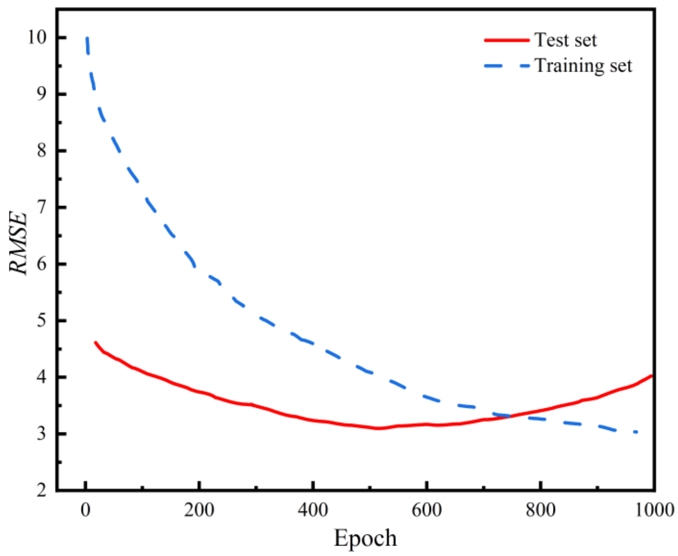


Fig 12. Model Prediction Error Curve During Incremental Learning.

The impact of training epochs on the prediction error of the torsional characteristics model for the driven disc assembly is shown in Figure 12. It can be observed that as the number of training epochs increases, the training loss continuously decreases, while the validation error initially decreases and then rises. This indicates that when the number of training epochs is less than 500, the model has not yet fully learned the features of the new data (underfitting); whereas when the epochs exceed 500, the model begins to overfit by learning noise features (overfitting). Based on this inflection point, the optimal number of training epochs is determined to be 500. After retraining with this setting, the RMSE on the test dataset reaches 3.045 N·m/°. After incorporating the incremental data, the optimized neural network hyperparameter configuration for the torsional stiffness prediction model—determined using the grid search method—is shown in Table 6.

Table 6. Parameter Settings of the ANN Model Optimized by Incremental Learning.

Hyperparameter Name		Preset Value
Number of Network Layers		3
Input Layer	Number of Neurons	6
	Activation Function	/
Intermediate Layer	Number of Neurons	50
	Activation Function	sigmoid
Output Layer	Number of Neurons	1
	Activation Function	Tanh Function
Loss Function		*RMSE*
Evaluation Metric		R^2
Number of Training Epochs		500
Batch Size		3

5.3. Model Optimization Results

During the incremental learning phase, the same network architecture as the initial artificial neural network model is adopted—namely, a feedforward neural network with two hidden layers. The number of neurons in each layer is set to 64 and 32, respectively, and the ReLU function is used as the activation function. To enhance the model's adaptability to new data while avoiding overfitting during training, several adjustments were made to the training parameters: the learning rate was set to 0.005, slightly lower than the initial model's 0.01, to reduce disruption to existing weights; the optimizer remains Adam, preserving its adaptive learning rate properties; and mini-batch training is employed, with a batch size of 3 per iteration. In addition, during the incremental learning process, the parameters of the first hidden layer were frozen, and only the weights of the second hidden layer and the output layer were updated. This strategy preserves the original model's ability to extract features from the old data while maintaining a balance between previously learned knowledge and newly acquired information.

The comparison of evaluation metrics for the three basic models (SVM, RF, and ANN) and the ANN optimized by incremental learning is shown in Table 7 below. The root mean square error (RMSE) of the torsional stiffness prediction model decreased from 5.471 N·m/° before incremental learning to 3.045 N·m/°, indicating a significant enhancement in predictive precision. The accuracy of the axial compression prediction model also improved, with the RMSE reaching 0.0104 mm—bringing the predicted results very close to the actual values. This demonstrates that, upon incorporating new training data, the ANN models achieved substantial performance gains.

Table 7. Performance Comparison of the ANN Model Optimized by Incremental Learning and the Three Basic Models.

Model	R^2	*RMSE*	CPU Time
SVM	0.9133	10.323	270ms
RF	0.9388	6.483	2.6s
ANN	0.9590	5.471	5.4s
ANN Model Optimized by Incremental Learning	0.9891	3.045	1.6s

As illustrated in Figure 13, the torsional stiffness prediction model achieved a maximum prediction error of 6.323 N·m/°, where the predicted value exceeded the actual value, and a minimum error of 0.457 N·m/°.

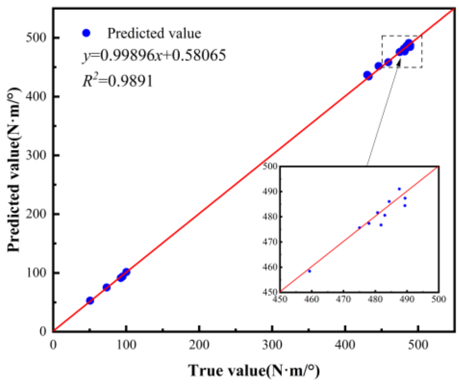


Fig 13. Torsional Stiffness Prediction Results of the ANN Model Optimized by Incremental Learning

In the preceding study, three machine learning models were applied to model the dataset from torsional characteristic tests of clutch driven disc assemblies, resulting in two optimal Artificial Neural Network (ANN) models. The R^2 and *RMSE* values of these models were 0.9590 and 5.471 N·m/°, and 0.9735 and 0.0161 mm, respectively. These two ANN models are hereafter referred to as the baseline models. Subsequently, incremental learning optimization was performed on the baseline models using newly acquired sample data.

After optimization, the updated models achieved improved performance on the test set, with R^2 and *RMSE* values of 0.9891 and 3.045 N·m/°, and 0.9932 and 0.0104 mm, respectively. To visually quantify the improvements from incremental learning, scatter plots comparing predicted versus actual values were generated for both the baseline and optimized models using 20 randomly selected driven disc assembly samples, as shown in Figure 14.

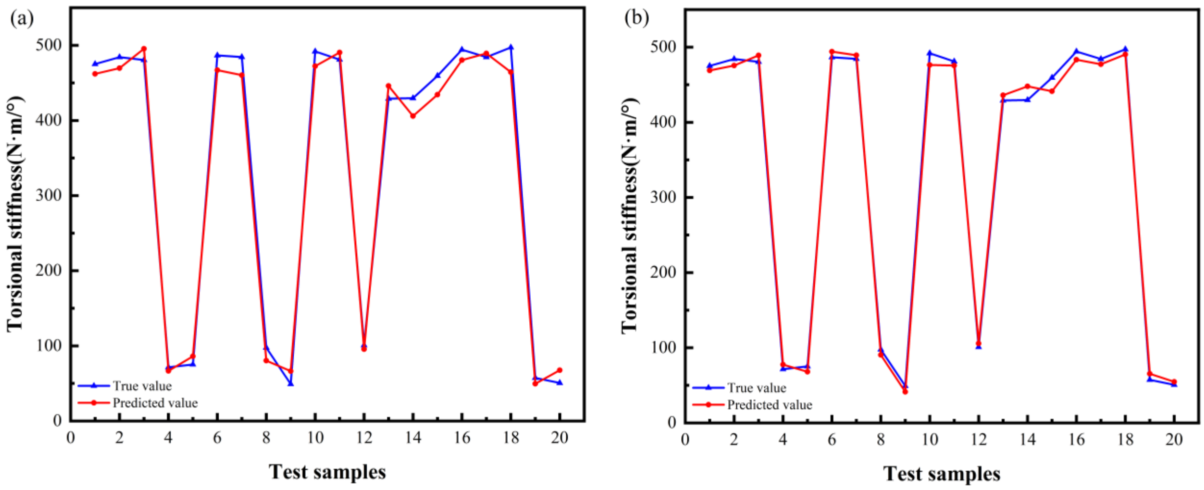


Fig 14. Comparison of Predicted and Actual Values of Torsional Stiffness: (a)ANN Model, (b)ANN Model Optimized by Incremental Learning.

Figure 14(a) and Figure 14(b) show the prediction comparisons on the test set between the original model and the incrementally optimized model for the torsional characteristics of the clutch driven disc assembly. As observed in Figure 14(b), the incrementally optimized model demonstrates a better fit to the test data compared to the original model. The optimized model not only exhibits reduced fitting errors but also, through incremental learning and adaptation to the new test data, more effectively captures the underlying features of the sample data.

In conclusion, incremental learning optimization significantly reduced the model's prediction error and improved its overall performance. This effect is particularly notable in the modeling of torsional characteristic test data for clutch driven disc assemblies, where increasing the volume of training samples effectively enhances model performance.

6. Discussion

This study focuses on predictive modeling of torsional stiffness in clutch driven disc assemblies, examining the development, optimization, and performance evaluation of machine learning algorithms, including SVM, RF, and ANN. By using production process parameters and torsional characteristic test data as inputs, with torsional stiffness as the output, we propose an ANN model integrated with incremental learning. This approach enables continuous updates of the model's knowledge structure as the training dataset (138 samples) expands, thereby enhancing both generalization and predictive accuracy. The optimized ANN model demonstrates superior performance, achieving an R^2 of 0.9891 and an *RMSE* of 3.045 N·m/°. Figure 2 presents the machine learning-based torsional stiffness prediction process, while Figure 11 illustrates the flowchart of the incremental learning algorithm. Collectively, this research provides a novel solution for predicting the torsional characteristics of clutch driven disc assemblies and offers valuable insights into the application of machine learning for engineering prediction.

In comparison with related studies, Georgiou et al. [13] employed 200 samples to predict the response characteristics of flush end-plate beam–column connections using ANNs, but their limited dataset raised concerns about prediction accuracy. Although our study utilized 138 samples, the integration of incremental learning effectively expanded the dataset and optimized the model, mitigating the impact of limited data. Hu et al. [14] applied RF, Decision Trees, and SVM to predict brake disc balancing errors; however, their reliance on traditional models without optimization strategies resulted in relatively low accuracy. By contrast, our incremental learning–optimized ANN model achieves higher predictive precision. Similarly, Zeng et al. [15] adopted an Extreme Learning Machine (ELM) for automotive engine torque prediction, highlighting the potential of single hidden layer neural networks, yet they did not incorporate optimization algorithms to further improve model performance.

Table 8 presents a comparative analysis between our study and previous research. However, since studies specifically focusing on the clutch driven disc assembly are very limited, we had to broaden the scope of comparison beyond the prediction of the driven disc assembly alone. Therefore, the emphasis of the comparison was placed on the predictive accuracy of the models.

Table 8. Comparative analysis of this study with the latest research in the field.

Study	Optimal model	R^2	*RMSE*
Proposed Model	ANN+ Incremental learning	0.9891	3.045
Aslan E [10]（2024）	K-Nearest Neighbor Regressor	0.9872	2.16
Alpsalaz F [11] (2025)	GBE	0.9972	0.2582
Hu Y et al. [14] (2023)	RF	-	0.5801
Zeng W et al. [15] (2020)	ELM+ANN	-	9.18
Özüpak Y et al. [18] (2025)	RNN+LSTM	0.9901	0.0187
Ansari A et al. [21]（2024）	XGBoost	0.9993	1.4744

As shown in Table 8, the proposed ANN combined with incremental learning achieves a high coefficient of determination (R^2=0.9891) and a relatively low prediction error (*RMSE* = 3.045), indicating strong predictive capability. In comparison, the ELM+ANN model by Zeng W et al. [15] (*RMSE* = 9.18) and the RF model by Hu Y et al. [14] (*RMSE* = 0.5801) perform significantly worse, while the KNN model by Aslan E [10] (R^2=0.9872) and the RNN+LSTM model by Özüpak Y et al. [18] (R^2=0.9901) achieve results comparable to the proposed method. Although the GBE model by Alpsalaz F [11] (R^2=0.9972, *RMSE* = 0.2582) and the XGBoost model by Ansari A et al. [21] (R^2=0.9993, *RMSE* = 1.4744) demonstrate superior overall performance, these approaches are typically trained in static settings and lack adaptability to dynamically updated data. In contrast, the proposed model, by incorporating an incremental learning mechanism, maintains stable predictive performance as the dataset gradually increases, thereby exhibiting stronger scalability and practical engineering value. Therefore, the proposed method not only ensures high prediction accuracy but also provides dynamic adaptability and practicality, offering a feasible and efficient solution for related engineering applications.

7. Conclusion

This study developed a comprehensive model training framework applicable to three types of machine learning algorithms. A systematic methodology was established, encompassing data preprocessing techniques and hyperparameter tuning strategies, and its effectiveness was thoroughly validated. Using torsional characteristic test data of clutch driven disc assemblies, the proposed framework was employed to train and optimize a torsional stiffness prediction model. Among the tested algorithms, the artificial neural network (ANN) achieved the best performance, with an R^2 of 0.9590 and an *RMSE* of 5.471 N·m/°.

In addition, an incremental learning procedure based on neural networks was designed. This process involves preprocessing and normalizing new data, loading the existing model, and applying an incremental learning function to enable online updates and dynamic model optimization. When 100 additional torsional test datasets were used as incremental input, the optimized ANN model achieved an R^2 of 0.9891 and an *RMSE* of 3.045 N·m/°, demonstrating significantly improved prediction accuracy and the ability to dynamically update the model.

Overall, this study demonstrates the strong potential of machine learning in predictive applications, extending beyond engineering to domains such as disease prediction and early diagnosis, stock market forecasting, traffic flow analysis, climate change modeling, crop yield estimation, and energy demand forecasting. These examples highlight the broad applicability of machine learning across diverse industries, offering transformative opportunities for many fields. In particular, the incremental learning approach proposed in this study significantly enhances the predictive capability of ANNs, outperforming unoptimized neural network models and showing strong promise for real-time applications. Looking ahead, future research could focus on integrating interpretability techniques such as SHAP or LIME to improve model transparency in industrial applications, as well as adopting efficient algorithms like XGBoost to achieve a better balance between accuracy and computational efficiency. Such efforts would further support real-time quality control in manufacturing environments and expand the scope of machine learning applications as data resources and computational technologies continue to advance.

Acknowledgements

The authors gratefully acknowledge the financial support from the Jilin Province Science and Technology Development Plan (No. 20220201027GX).

References

1. Yu L, Ma B, Chen M , Li H , Liu J & Li M. Investigation on the failure mechanism and safety mechanical-thermal boundary of a multi-disc clutch. Engineering Failure Analysis. 103, 319-334 (2019) http://doi.org/10.1016/j.engfailanal.2019.04.071
2. Shangguan W, Liu X, Rakheja S & Hou Q. Effective utilizing axial nonlinear characteristics of diaphragm spring and waveform plate to enhance breakaway performances of a clutch. Mechanical Systems and Signal Processing. 125, 123-141(2019) http://doi.org/10.1016/j.ymssp.2018.05.060
3. Kim S, Lee H, Kim J & Park G. Online adaptive identification of clutch torque transmissibility for the drivability consistency of high-performance production vehicles. Control Engineering Practice. 147, 105926(2024) http://doi.org/10.1016/J.CONENGPRAC.2024.105926
4. Wei Z, Shangguan W, Liu X & Hou Q. Modeling and analysis of friction clutches with three stages stiffness and damping for reducing gear rattles of unloaded gears at transmission. Journal of Sound and Vibration. 483, 115469(2020) http://doi.org/10.1016/j.jsv.2020.115469
5. Liu X, Shangguan W, Jing X & Ahmed Wet. Vibration isolation analysis of clutches based on trouble shooting of vehicle accelerating noise. Journal of Sound and Vibration. 382, 84-99(2016) http://doi.org/10.1016/j.jsv.2016.07.008
6. Zhang C, Yu W, Zhang Y, Xu J & Zeng Q. Dynamics modeling and analysis of the multistage planetary gear set-bearing-rotor-clutch coupling system considering the tooth impacts of clutches. Mechanical Systems and Signal Processing. 214, 111365 (2024) http://doi.org/10.1016/J.YMSSP.2024.111365
7. Abdullah IO and Schlattmann J. Thermal behavior of friction clutch disc based on uniform pressure and uniform wear assumptions. Friction. 4(3), 228-237(2016) http://doi.org/10.1007/s40544-016-0120-z
8. Neupert T, Benke E & Bartel D. Parameter study on the influence of a radial groove design on the drag torque of wet clutch discs in comparison with analytical models. Tribology International. 119, 809-821(2018) http://doi.org/10.1016/j.triboint.2017.12.005

9. Jiang Y, Ma J, Chen D, Liu Z L, Li Y & Paik J. Compact Pneumatic Clutch With Integrated Stiffness Variation and Position Feedback. IEEE Robotics and Automation Letters. 6(3), 5697-5704(2021) http://doi.org/10.1109/LRA.2021.3083236
10. Aslan E. Temperature Prediction and Performance Comparison of Permanent Magnet Synchronous Motors Using Different Machine Learning Techniques for Early Failure Detection. Eksploatacja i Niezawodnosc – Maintenance and Reliability. 27(1) (2024) http://doi.org/10.17531/EIN/192164
11. Alpsalaz F. Fault Detection in Power Transmission Lines: Comparison of Chirp-Z Algorithm and Machine Learning Based Prediction Models. Eksploatacja i Niezawodnosc – Maintenance and Reliability. (2025) http://doi.org/10.17531/ein/203949
12. Henriques L, Farinha T & Mendes M .Fault Detection and Prediction for a Wood Chip Screw Conveyor. Eksploatacja i Niezawodnosc – Maintenance and Reliability. 26(3) (2024) http://doi.org/10.17531/EIN/189323
13. Gregory G and Ahmed E. ANN-Based Model for Predicting the Nonlinear Response of Flush Endplate Connections. Journal of Structural Engineering. 150(5) (2024) http://doi.org/10.1061/JSENDH.STENG-13119
14. Hu Y, Lv W, Wang Z, Liu L & Liu H. Error prediction of balancing machine calibration based on machine learning method. Mechanical Systems and Signal Processing. 184(2023) http://doi.org/10.1016/J.YMSSP.2022.109736
15. Zeng W, Khalid S A M, Han X & Jimi T. A Study on Extreme Learning Machine for Gasoline Engine Torque Prediction. IEEE Access. 8, 104762-104774(2020) http://doi.org/10.1109/access.2020.3000152
16. Jang DW, Kang JS & Lim JY. A feasible strain-history extraction method using machine learning for the durability evaluation of automotive parts. Journal of Mechanical Science and Technology. 35(11), 5117-5125(2021) http://doi.org/10.1007/S12206-021-1028-8
17. Özüpak Y & Mansurov S. Optimizing electricity demand forecasting with a novel RNN-LSTM hybrid model. Energy Sources, Part B: Economics, Planning, and Policy. 20(1), (2025) https://doi.org/10.1080/15567249.2025.2531448
18. Özüpak Y, Alpsalaz F & Aslan E. Air Quality Forecasting Using Machine Learning: Comparative Analysis and Ensemble Strategies for Enhanced Prediction. Water Air Soil Pollut. 236, 464 (2025) https://doi.org/10.1007/s11270-025-08122-8
19. Aslan E, Özüpak Y, Alpsalaz F & Elbarbary Z M S, A Hybrid Machine Learning Approach for Predicting Power Transformer Failures Using Internet of Things-Based Monitoring and Explainable Artificial Intelligence. IEEE Access. 13, pp, 113618-113633(2025) https://doi.org/10.1109/ACCESS.2025.3583773
20. Özüpak Y, Alpsalaz F, Aslan E & Uzel H. Hybrid deep learning model for maize leaf disease classification with explainable AI. New Zealand Journal of Crop and Horticultural Science. 53(5). 2942–2964(2025) https://doi.org/10.1080/01140671.2025.2519570
21. Ansari A & Quaff R A .Advanced Machine Learning Techniques for Precise hourly Air Quality Index (AQI) Prediction in Azamgarh, India. International Journal of Environmental Research. 19(1). 15-15(2024) https://doi.org/10.1007/S41742-024-00684-5
22. Arrieta BA, Díaz-Rodríguez N, Ser D J, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R, Chatila R & Herrera F. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion. 58, 82-115(2020) http://doi.org/10.1016/j.inffus.2019.12.012
23. Davide C, J M W & Giuseppe J. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ. Computer science. 7, e623-e623(2021) http://doi.org/10.7717/PEERJ-CS.623
24. Nick P, Luis C & Francesco M. Adaptive learning for reliability analysis using Support Vector Machines. Reliability Engineering and System Safety. 226(2022) http://doi.org/10.1016/J.RESS.2022.108635
25. Atin R & Subrata C. Support vector machine in structural reliability analysis: A review. Reliability Engineering and System Safety. 233(2023) http://doi.org/10.1016/J.RESS.2023.109126
26. Sun J, Zhong G, Huang K & Dong J. Banzhaf random forests: Cooperative game theory based random forests with consistency. Neural Networks. 106: 20-29(2018) http://doi.org/10.1016/j.neunet.2018.06.006
27. Sagi O & Rokach L. Ensemble learning: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 8(4) (2018) http://doi.org/10.1002/widm.1249
28. Ghritlahre KH & Prasad KR. Application of ANN technique to predict the performance of solar collector systems - A review. Renewable and Sustainable Energy Reviews. 84, 75-88(2018) http://doi.org/10.1016/j.rser.2018.01.001
29. Abiodun I O, Jantan A, Omolara E A, Dada K V, Mohamed N A & Arshad H. State-of-the-art in artificial neural network applications: A survey. Heliyon. 4(11): e00938(2018) http://doi.org/10.1016/j.heliyon.2018.e00938

30. H. I S. Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective. SN Computer Science. 2(3) (2021) http://doi.org/10.1007/S42979-021-00535-6
31. Tufail S, Riggs H, Tariq M & Sarwat A I. Advancements and Challenges in Machine Learning: A Comprehensive Review of Models, Libraries, Applications, and Algorithms. Electronics. 12(8) (2023) http://doi.org/10.3390/ELECTRONICS12081789
32. Wang L, Zhang X, Su H & Zhu J. A Comprehensive Survey of Continual Learning: Theory, Method and Application. IEEE transactions on pattern analysis and machine intelligence. 8(46) (2024) http://doi.org/10.1109/TPAMI.2024.3367329
33. Liu BY & Liu JF. A Survey on Incremental Learning. Modern Computer 2022; 28(13): 72-75.
34. Han Y, Huang G, Song S, Le Y & Wang H. Dynamic Neural Networks: A Survey. IEEE transactions on pattern analysis and machine intelligence. 44(11), 7436-7456 (2022) http://doi.org/10.1109/TPAMI.2021.3117837
35. Liu H, Zhou Y, Liu B, Zhao J, Yao R & Shao Z. Incremental learning with neural networks for computer vision: a survey. Artificial Intelligence Review. 56(5), 4557-4589 (2022) http://doi.org/10.1007/S10462-022-10294-2