Chapter- 7

Results and Discussion

Analysis of Frictional properties of CNTs coated aramid FREC using ML Techniques

This chapter emphasize a computational and machine learning (ML)-based approach to analyze the tribological performance of CNT-coated aramid fabric-reinforced epoxy composites. The effects of mechanical properties, thermal characteristics, and tribological parameters (sliding frequency, normal load, and temperature) on the coefficient of friction (COF) were investigated. Predictive models were developed using Artificial Neural Networks (ANN), Gradient Boosting Machines (GBM), and Random Forest (RF) based on prior experimental data. These three models ANN, RF, and GBM were chosen because they offer complementary strengths in capturing complex patterns, handling diverse data types, and ensuring robust performance. ANN excels at learning nonlinear relationships and deep patterns in data. Random Forest provides high accuracy, resistance to overfitting, and works well with small datasets. GBM offers strong predictive power by sequentially improving errors, making it ideal for fine-tuned performance. Together, they balance interpretability, accuracy, and generalization. The feature importance chart is an efficient way to condense the impact of each independent material property and operational variable on forecasting the COF output. This data-driven ML analysis offers significant insights into the tribological behavior of fiber-reinforced polymer composites, aiding in material design and performance optimization.

7.1 Machine learning models

In this study a machine learning (ML) based data-driven technique was used. This strategy included pre-processing, optimization, and data collecting to increase overall performance [269]. The authors got started on an organized expedition to build a very precise version for anticipating the COF, a vital criterion in economic preparation. In this study, the prediction of the COF using ML models was executed through a comprehensive methodology that encompassed data collection, preprocessing, model training, and evaluation. Initially, a detailed dataset was compiled from previous experiments, incorporating various parameters that influence COF, such as CNTs content, applied load, sliding frequency, temperature, tensile strength, Vickers hardness, and thermal conductivity, with COF as the output variable. We utilized k-fold cross-validation to maximize the utility of our available data. By dividing the dataset into k subsets and training the model k times, each time using a different subset as the validation set and the remaining as the training set, we ensured that our models were thoroughly evaluated [270]. Three distinct ML algorithms were employed: the ANN, which utilized a deep neural network architecture with three hidden layers and 20 neurons per layer, applying the "relu" activation function and a learning rate of 0.022 to optimize prediction accuracy; the RF model was optimized by utilizing 100 decision trees (n-estimators) and 42 measured features (max-features). Which built multiple decision trees based on random subsets of the training data and averaged their predictions to enhance accuracy and mitigate overfitting; and the GBM, which trained sequentially to correct errors from previous trees, utilizing a learning rate of 0.01 and 100 trees (n-estimators) for optimal performance. The models were evaluated using metrics including MAE, MSE, RMSE, and R^2. The models COF predictions were compared with experimental data, demonstrating their effectiveness in modeling tribological behavior and providing valuable insights for material design and optimization. The whole design growth procedure was thoroughly carried out in Python within the Jupyter Notebook Editor, leveraging noticeable collections such as "Pandas" for information adjustment coupled with "scikit-learn" for mathematical execution and also evaluation. With this extensive method, the writers intended not just to create a precise COF forecast version but likewise to add beneficial understandings right into the usefulness together with the effectiveness of various artificial intelligence methods in monetary projecting [271].

7.1.1 Artificial neural network (ANN)

ANN works like a nonlinear signal-processing system, where interconnected neurons adjust weights through backpropagation and gradient descent to minimize prediction error, similar to minimizing energy in physical systems. ANNs are advanced tools that simulate the complex functions of the human mind. Consequently, ANNs are extremely reliable at translating fancy patterns along with partnerships discovered in datasets [272]. In this present research, an anticipated version for a very exact quote of the COF utilizing a specific type of ANN, referred to as a Deep Neural Network (DNN). Scaling the training datasets was an integral part of creating the ANN version. By transforming information worth's right into a constant array, this treatment quits any kind of one facet from unjustly affecting an additional as a result of variants in mathematical size. Our information was scaled, making use of the "MinMaxScaler" bundle, which transformed all worth to drop between 0 as well as 1. The input, concealed, and also outcome layers of the ANN style are comprised of distinct nerve cells and activation features that cooperate with a procedure and generate outcome information [273]. The resultant data then serves as input for the subsequent neuron, forming a sequential flow within the network. These neurons are arranged into columns that are called layers. Three layers make up the network: the input layer, hidden layers, and output layer. Typically, there are multiple hidden layers within the network, forming an extensive network. We established a Multilayer Perceptron (MLP) version for our examination, with 3 concealed layers and also 20 nerve cells in each. The feed-forward network of the MLP regressor architecture of the ANN is shown in Figure 7.1. We utilized the "relu" activation feature along with a small regularization term (a = 0.022) to make the most of the forecast efficiency of the version. Overfitting is a common trouble with ANN designs where the design succeeds in training information yet inadequately on fresh, unidentified information is decreased with normalization. Due to its calculating performance plus simplexes, the "relu" activation function, which implies the limited straight system is a recommended alternative for deep understanding designs. In Figure 7.1, input parameters X_1 to X_7 include CNTs content (%), load (N), frequency (Hz), temperature (°C), tensile strength (MPa), Vickers hardness (HV), and thermal conductivity (W/m·K), selected for their relevance to the tribological performance of the composite. These features were selected due to their significant influence on the frictional behavior and overall tribological performance of the composite material system.

Figure 7.2 presents heatmaps of Spearman's rank correlation coefficients to analyze the monotonic relationships between input features and target variables. The coefficients (R), ranging from -1 to 1, indicate the strength of the correlation: $|R| > 0.75$ represents a strong correlation, $|R| = 0$ indicates no correlation, and $|R| = 1$ implies full correlation. The input features exhibit no multicollinearity ($|R| = 0$), ensuring their independence and suitability for regression modeling. The COF correlation map indicates a strong correlation with load ($|R| = 0.78$), moderate correlations with frequency ($|R| = 0.66$) and materials ($|R| = 0.48$), and the weakest correlation with thermal conductivity ($|R| = 0.12$) among the input features. These findings validate the selection of input features for subsequent modeling.

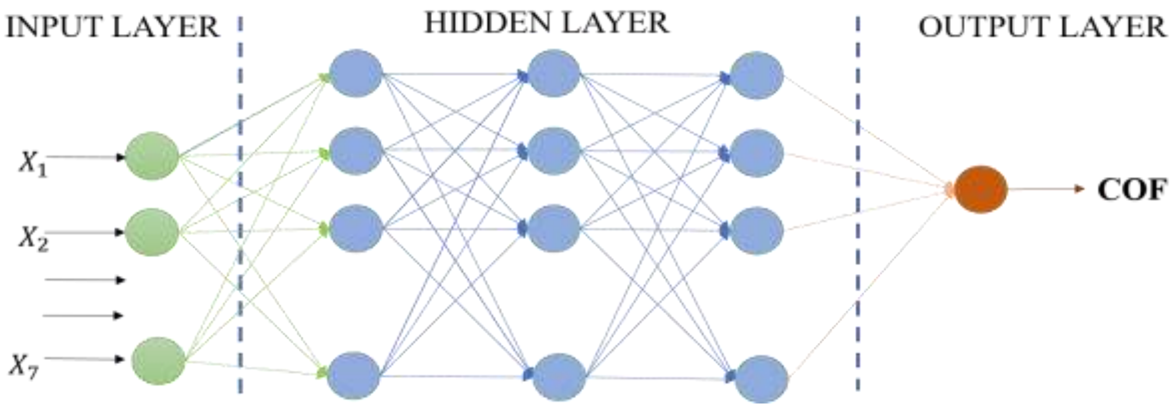

Figure 7.1 The architecture of the artificial neural network (ANN) consists of a three-layer perceptron with an input layer, hidden layers, and an output layer.

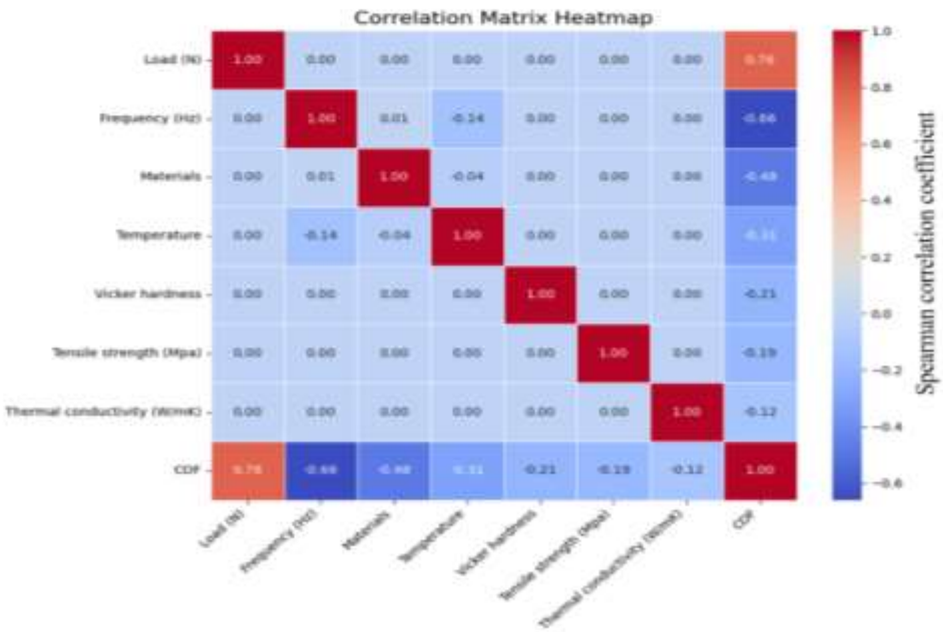

Figure 7.2 Spearman rank's feature correlation coefficient matrix heatmaps for COF

7.1.2 Random Forest (RF)

Random Forest models function as an ensemble of independent decision trees, each trained on a different random subset of the data, and their predictions are aggregated to produce a more robust and stable final result [274]. Random Forest is a powerful supervised machine learning algorithm that constructs numerous choice trees to enhance precision as well as take care of complex, nonlinear datasets with a great deal of inputs as shown in Figure 7.3. Various nodes refine information at each degree of the choice tree coupled with creating typical worth's for the degrees that comply with them. The RF version makes use of a bagging method in which a number of parts of the training information are made, and also choice trees are created for every part. This approach boosts the design's efficiency and also reduces overfitting, specifically with non-linear plus discontinuous datasets [275]. With the RF design, every choice tree is included in the last projection, just like in Freedom. Distinguishing the criteria together with the training information substratum of each tree makes it possible for a forecast that is a lot more accurate coupled with resistance than that of a single-controlled choice tree. Specific choice trees in the RF regressor version are based upon collections of regulations as well as features that use entropy and also information gain, two mathematical ideas. Information pureness can be gauged by making use of entropy, a mathematical procedure of information changeability. The RF design can determine one of the most essential functions to divide by determining entropy, making it possible for much better decision-making. The logarithm of the variety of courses coupled with the chance of each course is increased to establish entropy using this formula. This approach makes it feasible to gauge information randomness specifically, which is essential for developing trustworthy choice trees. The entropy calculation formula is given below [276]:

$$\text{Entropy (E(s))} = -\sum p\,(Z) * \log_2 p(Z) \tag{7.1}$$

where,

$p(Z)$ = probability of obtaining Z outcome classes.

The formula for information gain calculation is given below:

$$\text{Information Gain (IG)} = Entropy_{parent} - \sum Probality\,(Z_x) * Entropy_{child} \tag{7.2}$$

In the above context, information gain is computed at each node using the following formulae:

$$IG(Z, Y) = ENTROPY(Z) - \sum_{x \in Y} \frac{|Z_x|}{Z} * ENTROPY(Z_x) \quad (7.3)$$

Where Z = target, Y = column variable, and x = each child value.

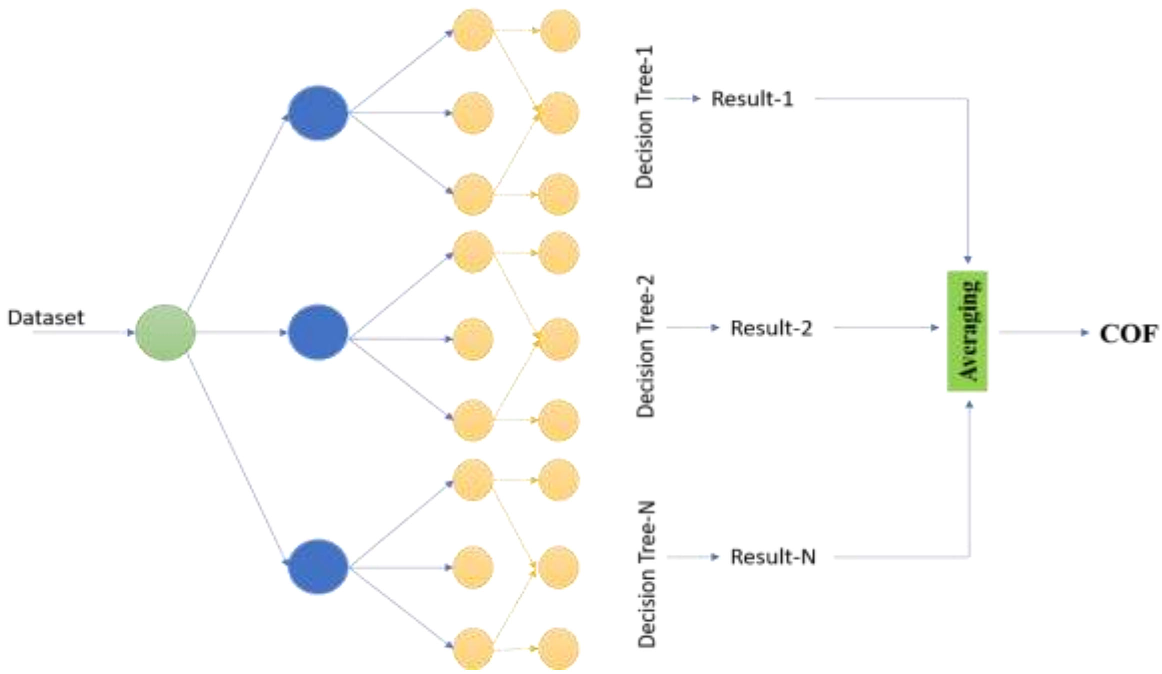

Figure 7.3 Schematic architecture of the RF model, featuring decision trees and a voting process.

For the Random Forest (RF) model, optimal performance in predicting COF was achieved by utilizing 100 decision trees (n-estimators) and 42 measured features (max-features) as key optimization parameters. In choice tree formulas, details gain is a crucial fact that develops the very best variable to split at each tree degree. Fundamentally, it determines just how much detail is removed from the information by sectioning it according to a details variable [277]. The variable utilized as the splitting requirement is the one that makes the most of info gain while decreasing entropy an action of information randomness. By making sure that every division in the choice tree makes the best use of the homogeneity of the occurring subgroups, this option treatment generates forecasts that are a lot more exact plus efficient. Choice trees can develop well-informed and also data-driven divides by offering concern to elements that have the best effect on decreasing unpredictability in the dataset which eventually boosts the version's efficiency.

7.1.3 Gradient boosting machine (GBM)

GBM functions as a sequential feedback-based control system, where each model corrects the previous one's errors using gradient descent, progressively reducing the system's overall loss like minimizing free energy step by step. Comparable to Random Forest (RF) the GBM version is an enhanced model that specifies its intricacy by making use of essential specifications like finding out price plus improving stages. Unlike RF, GBM utilizes an enhancing approach in which choice trees are constructed one after the other to minimize losses or mistakes. As seen in Figure 7.4, the set onward method is related to producing a series of regression trees that are built one after the other, making use of every one of the dataset's functions. In Figure 7.4, the range for error is (0 to 100) %, and the range for Iteration is (0 to 100). To develop much more precise projections, the information from each tree is combined after that. GBM offers additional benefits of optimization as well as step-by-step mistake improvement over RF. To offset blunders made in earlier cycles, loss features are maximized for each and also a tree, making use of the improving procedure. The finding out price, stood for by α, is an essential consider GBM that influences the accuracy as well as computational efficiency of the design. Greater precision is accomplished at a reduced understanding price; however, the models rise in calculating time [278].

$$\text{Computation time} \propto 1/\alpha \qquad (7.4)$$

where α = learning rate.

Additionally, by using arbitrarily chosen sub-training established at each model and also reducing the opportunity of overfitting, randomization throughout design installation boosts anticipating precision [279]. Using this technique, the GBM design is ensured to be trusted, solid, and also able to procedure detailed datasets as well as generate specific forecasts.

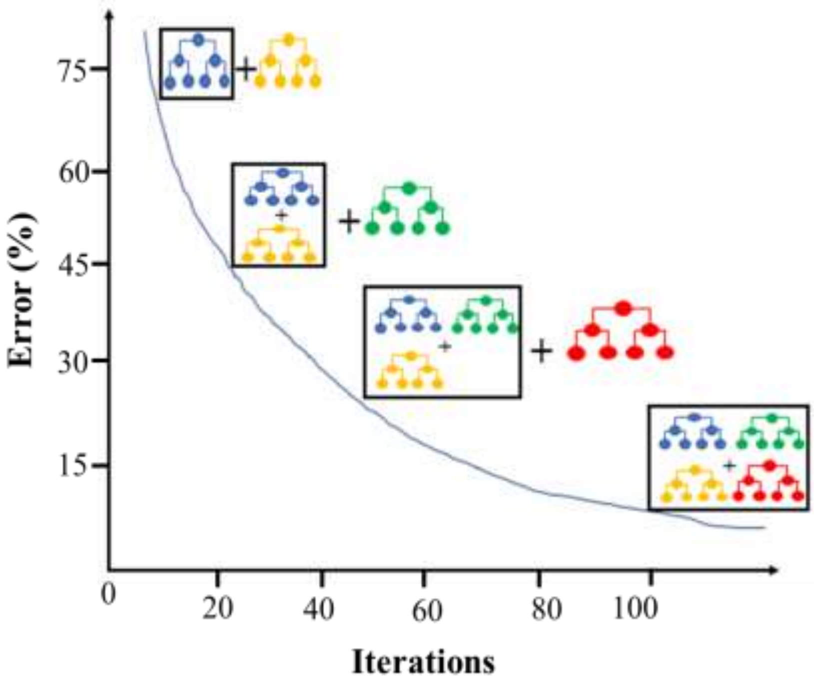

Figure 7.4 Algorithm for minimizing the loss function in gradient boosting machines.

7.2 Collection of data

Thorough data collection was achieved by carefully carrying out the experiment in a variety of settings and at different levels. Every effort was made to ensure precision and minimize errors, thereby maximizing accuracy. Consistency was maintained in each experiment, with data collected based on a singular experimental setup and consistent environmental conditions. This approach was adopted due to the significant influence of operating environmental conditions, experimental setup, and parameters on the COF. Utilizing advanced ML techniques, datasets were trained to discern trends and develop predictive models, thereby improving the prospects of achieving results with high accuracy. Data on AF/Epoxy and CAF-FCNT/Epoxy composites subjected to reciprocating wear under various circumstances were gathered throughout the experimental phase from earlier comparable tests carried out by the current researchers in Chapter 4. Particular information was gathered on the following topics: tribological operating factors (such as applied load, sliding frequency, temperature, and CNT content) and material qualities (such as tensile strength, thermal conductivity, Vickers hardness, etc.). As seen in Table 7.1, the Coefficient of friction (COF) was predicted using seven parameters and twenty datasets, for a total of 140 data points (20 x 7). Seven input variables, namely CNTs content, applied load, sliding frequency, temperature, tensile strength, Vickers hardness, and thermal conductivity, were examined. The COF served as the output parameter for the developed ML models.

7.3 Data preprocessing

Information preprocessing is crucial since well-prepared information is necessary to the efficiency of artificial intelligence versions. This involves numerous tasks, consisting of slicing the dataset right into training and also screening substratum to decrease predisposition, and identifying plus fixing abnormalities in the information. Numerical tools and computational libraries played a crucial role in data preprocessing, model development, training, and evaluation. Python was used as the primary programming language due to its simplicity, flexibility, and rich ecosystem of machine learning libraries. Scikit-learn was employed for implementing traditional models like Random Forest (RF) and Gradient Boosting Machine (GBM), offering efficient tools for model training, validation, and performance assessment. For ANN, MATLAB was utilized, providing a high-level interface for designing and training deep learning models. Additionally, NumPy and Pandas were used for numerical computations and structured data manipulation. These numerical tools ensured the accuracy, scalability, and reproducibility of the machine learning workflow, enabling effective modeling of the tribological data despite the limited dataset.

Table 7.1 Data collected for the purpose of creating an ML model for polymer composites from Chapter 4.

CNTs content %	Load (N)	Frequency (Hz)	Temperature (^{o}C)	Tensile strength (MPa)	Vickers hardness	Thermal conductivity (W/mK)	COF
0	30	8	30	250.79	21.42	0.155	0.166
0	40	8	30	250.79	21.42	0.155	0.191
0	50	8	30	250.79	21.42	0.155	0.203
0	60	8	30	250.79	21.42	0.155	0.227
0	40	6	30	250.79	21.42	0.155	0.236
0	40	10	30	250.79	21.42	0.155	0.175
0	40	12	30	250.79	21.42	0.155	0.164
0	40	8	40	250.79	21.42	0.155	0.183
0	40	8	50	250.79	21.42	0.155	0.173
0	40	8	60	250.79	21.42	0.155	0.159
1	30	8	30	310.063	28.06	0.21167	0.152
1	40	8	30	310.063	28.06	0.21167	0.185
1	50	8	30	310.063	28.06	0.21167	0.194
1	60	8	30	310.063	28.06	0.21167	0.21
1	40	6	30	310.063	28.06	0.21167	0.219
1	40	10	30	310.063	28.06	0.21167	0.169
1	40	12	30	310.063	28.06	0.21167	0.148
1	40	8	40	310.063	28.06	0.21167	0.179
1	40	8	50	310.063	28.06	0.21167	0.165
1	40	8	60	310.063	28.06	0.21167	0.154

We utilized k-fold cross-validation to maximize the utility of our available data. By dividing the dataset into k subsets and training the model k times, each time using a different subset as the validation set and the remaining as the training set, we ensured that our models were thoroughly evaluated. This circulation ensures a solid assessment of the design's capability for generalization. Furthermore, utilizing techniques such as

"MinMaxScaler" assists in just a range of information attributes between 0 as well as 1, which enhances convergence throughout version training. Such pre-processing methods can significantly enhance the anticipated precision plus reliability of the design.

7.4 Parameter optimization

There are several strategies for identifying the optimal criteria such as Gradient-Based Optimization, Grid Search as well as Bayesian Optimization. Yet, in this job, the researchers took a systematic method, making use of Python code to check out different criterion arrays. Throughout this treatment, an Artificial Neural Network (ANN) version with 3 hidden layers, each consisting of 20 neurons with a regularization term of 0.022 plus the "relu" activation feature, was picked to forecast the COF. Durable criterion choice was made by using grid search as well as cross-validation techniques to attain this option. Table 7.2 offers the very carefully recorded enhanced criteria for numerous artificial intelligence regression designs, consisting of ANN, RE, and GBM supplying informative details for efficiency analysis and also design enhancement.

Table 7.2 Optimization of COF models.

Model name	Optimization parameters
ANN	hidden layer (20,20,20), alpha= 0.022, activation function: “relu”
RF	maximum feature 42, n-estimators= 100
GBM	learning rate= 0.01, n-estimators= 100, max depth= 7

7.5 Model Performance Evaluation

Important metrics for assessing model performance include the root mean square error (RMSE), mean absolute error (MAE), mean squared error (MSE), and coefficient of determination (R^2 values) [164, 280]. Table 7.3 provides specific performance metrics for ML models that predict COF for epoxy composites reinforced with aramid fibers. The combination of these error metrics confirms a high degree of accuracy in the data-driven models. Table 7.3 compares the predicted COF of polymer composites using ANN, RF, and GBM algorithms against actual values. The R^2 value is a dominant factor in determining model performance. The R^2 has a range of 0 to 1, with 0 denoting no correlation and the model's inability to anticipate variations in the data. An R^2 less than 0.5 signifies a nominal correlation, implying the model's inadequacy for data prediction. An R^2 between 0.7 and 0.9 denotes an acceptable model, while an R^2 over 0.9 denotes an excellent model fit. The R^2 values for the ANN, RF, and GBM models were, respectively, 0.9088, 0.85294, and 0.92807. Table 7.3 demonstrates that the GBM model performed the best, with an R^2 of 0.92807, an MAE of 0.00335, an MSE of 0.0000165, and an RMSE of 0.00406. This model accurately predicted COF with a maximum accuracy of 92.807%. The RF model using a learning rate of 0.01, n-estimators= 100, and max_depth= 7 yielded the optimal prediction performance for COF. Consistent performance was also guaranteed by this approach when categorical variables were present in the COF data. With an R^2 of 0.9088, MAE of 0.00432, MSE of 0.0000907, and RMSE of 0.00527, the ANN model also demonstrated excellent performance on the intricate COF dataset. The generated ML models can reliably predict the COF of polymer composites based on mechanical and tribological test factors, depending on these performance criteria.

Table 7.3 Performance metrics of COF prediction models

ML model	MAE	MSE	RMSE	R^2 value
ANN	0.00432	0.00009	0.00527	0.9088
RF	0.0059	0.00042	0.00652	0.85294
GBM	0.00335	0.00002	0.00406	0.92807

Table 7.4 Comparison of the predicted COF by the ANN, RF, and GBM models with the experimental COF data.

COF	ANN predicted	ANN error	RF predicted	RF error	GBM predicted	GBM error
0.166	0.1589	0.00706	0.1751	-0.00909	0.1642	0.00184
0.191	0.1909	0.00001	0.1890	0.00196	0.1852	0.00573
0.203	0.2168	-0.01386	0.1972	0.00578	0.2039	-0.00093
0.227	0.2269	0.00002	0.2189	0.00812	0.2258	0.00116
0.236	0.2039	0.03202	0.2261	0.00985	0.2344	0.00162
0.175	0.1749	0.00002	0.1764	-0.0014	0.1754	-0.00036
0.164	0.1639	0.00004	0.1641	-0.0001	0.1637	0.00031
0.183	0.1835	-0.00052	0.1845	-0.00155	0.1832	-0.00018
0.173	0.1729	0.00005	0.1744	-0.00146	0.1659	0.00706
0.159	0.1589	0.00007	0.1661	-0.00717	0.1588	0.00015
0.152	0.1521	-0.00010	0.1654	-0.01346	0.1529	-0.00095
0.185	0.1850	-0.00004	0.1818	0.00322	0.1826	0.00242
0.194	0.1939	0.00007	0.1894	0.00457	0.1945	-0.00055
0.21	0.1973	0.01267	0.2101	-0.00007	0.2130	-0.00305
0.219	0.2035	0.01547	0.2168	0.00223	0.2193	-0.00033
0.169	0.1689	0.00001	0.1713	-0.00227	0.1695	-0.00055
0.148	0.1479	0.00002	0.1543	-0.00633	0.1494	-0.00146
0.179	0.1808	-0.00179	0.1791	-0.00010	0.1808	-0.00185
0.165	0.1651	-0.00017	0.1679	-0.00298	0.1648	0.00018
0.154	0.1566	-0.00263	0.1589	-0.00488	0.1551	-0.00109

The actual experimental data and the numerical predictions produced by the models were compared in order to assess the effectiveness of the machine learning (ML) models for forecasting the COF of the three composites. The predicted and experimental COF for aramid fiber-reinforced epoxy composites are compared in Table 7.4 utilizing ANN, RF, and GBM algorithms; this comparison is also visually shown in Figure 7.5(a)-(c). The implementation of machine learning algorithms for the production of components composed of polymer composite materials is validated by the small difference found between the anticipated values and the experimental findings. It can be seen from Table 7.4 that most data sets have errors that are quite near to zero. The accuracy of the forecasts was also assessed by utilizing the absolute percentage error of the models. Based on the experimental data and anticipated COF outputs, Table 7.5 displays the absolute error and absolute error percentage. It shows that the absolute error for most data sets falls between 0.00001 and 0.0320. As shown in Table 7.5, the ANN model exhibited the highest error percentage (13.56%) and also demonstrated the lowest error percentage (0.003658%) among the three models for certain test sets. The smallest error percentages during this study were 0.003658%, 0.03214%, and 0.095624% for the ANN, RF, and GBM models, respectively. According to Table 7.3, the GBM model achieved the best prediction performance, with values closely matching the actual COF. Although all three models predicted the COF satisfactorily, the GBM and ANN models provided the most accurate results.

Table 7.5 Experimental data and absolute error percentage

COF	ANN error	RF error	GBM error	ANN error %	RF error %	GBM error %
0.166	0.00706	0.00909	0.00184	4.2564	5.4744	1.1110
0.191	0.00001	0.00196	0.00573	0.0036	1.0262	2.9994
0.203	0.01386	0.00578	0.00093	6.8264	2.8461	0.4562
0.227	0.00002	0.00812	0.00116	0.0071	3.5793	0.5097
0.236	0.03202	0.00985	0.00162	13.5662	4.1748	0.6844
0.175	0.00002	0.0014	0.00036	0.0109	0.7743	0.2046
0.164	0.00004	0.0001	0.00031	0.0215	0.0625	0.1874
0.183	0.00052	0.00155	0.00018	0.2827	0.8484	0.0987
0.173	0.00005	0.00146	0.00706	0.0318	0.8410	4.0805
0.159	0.00007	0.00717	0.00015	0.0438	4.5110	0.0956
0.152	0.00010	0.01346	0.00095	0.0676	8.8536	0.6218
0.185	0.00004	0.00322	0.00242	0.0201	1.7432	1.3085
0.194	0.00007	0.00457	0.00055	0.0377	2.3570	0.2813
0.21	0.01267	0.00007	0.00305	6.0352	0.0321	1.4503
0.219	0.01547	0.00223	0.00033	7.0626	1.0205	0.1525
0.169	0.00001	0.00227	0.00055	0.0046	1.3446	0.3233
0.148	0.00002	0.00633	0.00146	0.0154	4.2787	0.9831
0.179	0.00179	0.00010	0.00185	1.0043	0.0545	1.0354
0.165	0.00017	0.00298	0.00018	0.1054	1.8091	0.1081
0.154	0.00263	0.00488	0.00109	1.7089	3.1656	0.7052

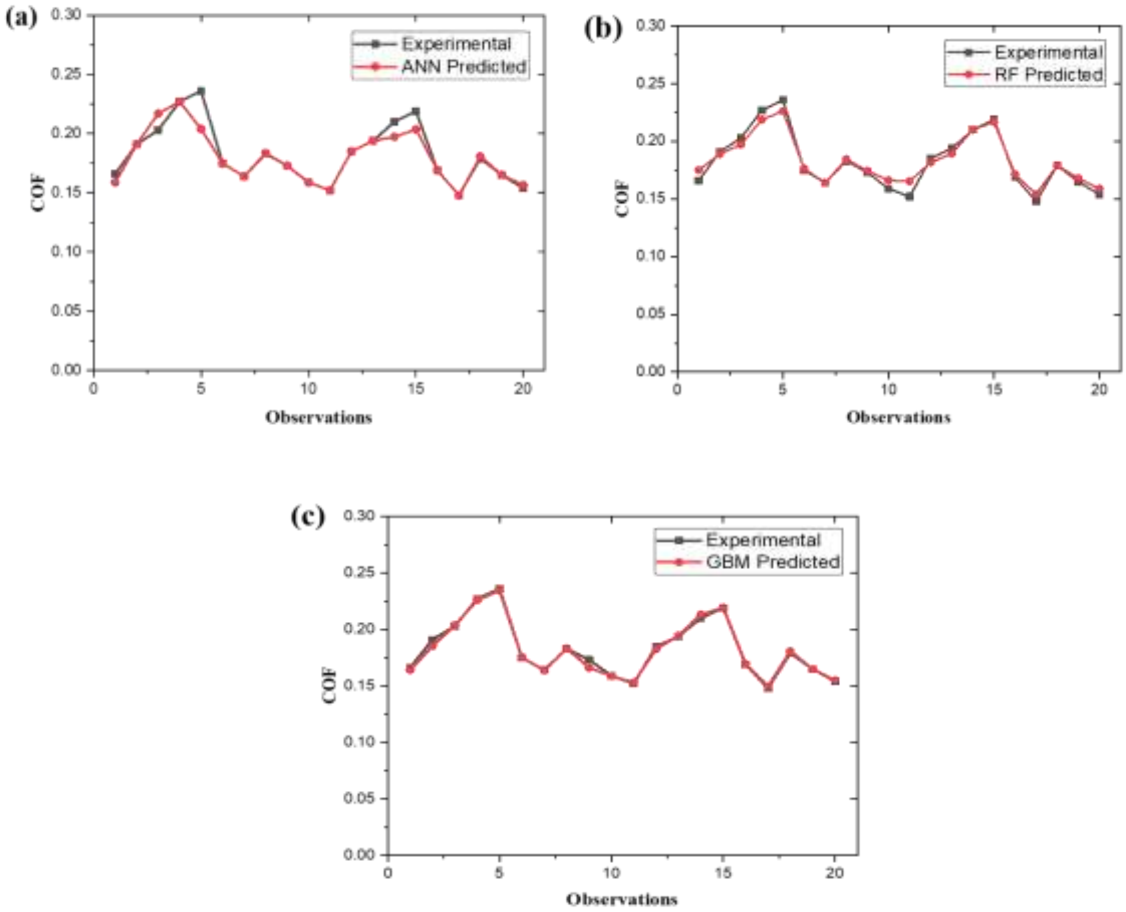

Figure 7.5 AF/Epoxy and CAF-FCNT/Epoxy composites' predicted COF data compared to experimental data using the (a) ANN, (b) RF, and (c) GBM models, respectively.

7.6 Feature importance for predicting the Coefficient of friction

The feature importance chart produced by decision tree-based machine learning models, like Random Forest (RF) and Gradient Boosting Machine (GBM), is an efficient way to condense the impact of each independent material property and operational variable on forecasting the COF output. Each independent variable in this chart is given a value between 0 and 1, with a total score of 1 representing the sum of all the variables' scores. An individual score of zero suggests no contribution to the output forecast, but a higher score suggests a more substantial impact. The fact that each input variable has a score that is not zero indicates that each of the variables that were chosen has an effect on predicting the coefficient of friction. Figures 7.6(a) and (b) present the feature importance charts generated by the GBM and RF algorithms for predicting the COF, and it was observed that the load, sliding frequency, and CNTs contents were the key factors in determining the COF prediction.

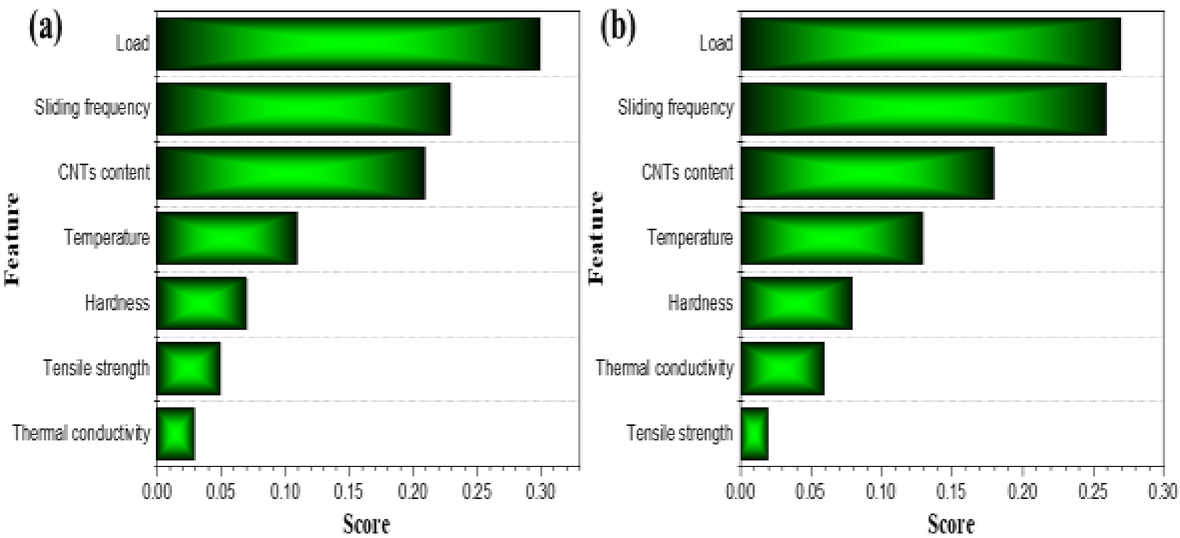

Figure 7.6 Using the models (a) GBM and (b) RF, the relative relevance of input variables for predicting the COF is based on feature importance.

According to the feature significance charts, load is the operational parameter that has the biggest impact on COF prediction since it has the highest score in both the RF and GBM models. An increase in contact pressure between the specimen and the steel ball under higher loads may lead to a rise in interfacial temperature, thereby enhancing the adhesion component of friction [217]. Additionally, the abrasion component of the friction coefficient intensifies significantly at elevated loads due to deeper penetration of asperities, resulting in higher frictional forces. These phenomena collectively provide a plausible explanation for the observed increase in the coefficient of friction under higher load conditions. The creation and maintenance of the lubricating film are directly related to the interaction between asperities on the sliding surfaces and the usual load. Furthermore, the inclusion of CNTs is essential for improving the self-lubrication effect and lowering friction. The self-lubricating properties of CNTs contribute to forming a carbon film on contacting surfaces, effectively reducing both friction and wear rates [213]. This effect is further enhanced by the rod-like structure of CNTs, which not only toughens polymer composites but also provides superior lubrication to worn surfaces [215]. This reduction is attributed to the decreased coefficients of adhesion and deformation, as well as improved interfacial adhesion between CNT-coated aramid fibers and the matrix, which collectively enhance tribological performance [216]. As initially highlighted, tribological properties are not intrinsic to the material but are influenced by various external and internal conditions. According to Unal et al. [281], materials experience wear through different mechanisms, which are contingent upon the material properties, environmental factors, and operating conditions. Numerous researchers have identified various factors influencing the frictional performance of fiber-reinforced polymer composites. Employing data-driven machine learning (ML) analysis allows for a comparative evaluation of various mechanical behavior and tribological operating variables, provide further insight into the tribological properties of polymer composites [282-284].

7.7 Summary

The finding can be summarized as follows:

- Machine learning techniques were employed to analyze the coefficient of friction (COF) in aramid fiber-reinforced epoxy composites by evaluating critical parameters, including CNTs content, load, sliding frequency, temperature, tensile strength, Vickers hardness, and thermal conductivity.
- Three ML models—Artificial Neural Networks (ANN), Random Forest (RF), and Gradient Boosting Machine (GBM)—were trained and evaluated.

- GBM achieved the highest accuracy ($R^2 = 0.92807$) with minimal error (MAE = 0.00335, RMSE = 0.00406), outperforming ANN ($R^2 = 0.9088$) and RF ($R^2 = 0.85294$).
- Feature importance analysis identified load, sliding frequency, and CNTs content as the most influential factors for COF.
- The ML-driven approach enhances the understanding of fiber-reinforced polymer composites, aiding in material design and optimization.