Towards an Explainable Multi-Target Regression, for Wear and Friction Prediction for Brake Pad Materials

Amira Sellami[1,2,*], Mouna Rekik[2,3], Chakib Ben Njima[2] and Riadh Elleuch[1]

[1]Laboratory of Electro-Mechanical System (LASEM), National School of Engineers of Sfax (ENIS), University of Sfax, Tunisia

[2]University of Sousse, Higher Institute of Transport and Logistics of Sousse, University of Sousse, Tunisia

[3]Laboratory of Multimedia, Information systems and Advanced Computing, Tunisia

E-mails: amira.sellami@istls.u-sousse.tn; mouna.rekik@istls.u-sousse.tn; chakib.bennjima@enim.rnu.tn; riadh.elleuch@gnet.tn

*Corresponding author

Abstract: The primary objective of this study is to create an effective multi-target regression model able to predict friction coefficient and wear rate, which are critical parameters for the tribological performance of brake systems. Two models, namely Random Forest (RF) and eXtreme Gradient Boosting (XG), were evaluated using performance metrics such as mean squared error, mean absolute error, and R-squared. In comparing to 1.2, 0.567, 0.59 for RF algorithm, the XG algorithm proves to be the more accurate model with MSE, MAE and R-squared respectively equal to 0.857, 0.4138, 0.756. XG (Extreme Gradient Boosting) outperforms RF (Random Forest) in terms of predictive accuracy in the specified prediction scenario, and the predicted results show good concordance with real values. However, a notable challenge with this model is the lack of interpretability, often referred to as a "black-box." In response to this issue, the study offers a comprehensive explanation, regarding as to how the XG model learns. Shapely Additive explanation model demonstrates that sliding speed is the most influential factor, positively affecting friction coefficient and wear rate of brake pad materials. In summary, the study contributes to the development of a machine learning model, that is accurate and explainable for the prediction of tribological performance in the field of brake pad materials.

Keywords: Extreme Gradient Boosting; Multitarget regression; Random Forest; Tribological performance; brake pad materials

1 Introduction

Optimizing the performance and safety of braking systems is a major concern in the automotive industry and many other industries where braking systems are crucial to ensure safe operations. The brake pad plays a crucial role in the braking system and is responsible for causing vehicle deceleration and stopping. The brake performance of a vehicle is conditioned by the tribological behavior of brake lining materials [1]. Various complex physical phenomena occur during braking at the pad/disc interface, leading to the formation of a third body, which is mainly affected by braking conditions: normal load, sliding speed [2], frictional heat, and material composition [3]. Therefore, it has a great impact on braking performance, especially braking distance and frictional heat generated at the pad/disc interface. Wear rate and friction coefficient are both key indicators of the tribological behavior of brake pad materials [4]. The complexity of contact phenomena [5] is certainly the reason why friction and wear are still so difficult to quantify. As a result, we frequently encounter the necessity to estimate these two critical factors that play a vital role in designing brake systems. In fact, the friction coefficient determines braking force, and the wear rate influences the durability of the brake pad material. Therefore, predicting the tribological performance of brake lining materials is imperative to improve their durability, efficiency and safety.

In general, simulating the material properties involves developing a model derived from experimental data. The challenge lies in the absence of effective prediction methods, stemming from difficulties in leveraging experimental data to accurately describe and comprehend the complex behavior of brake friction materials, as well as capturing the influence of material properties and test parameters on the wear rate. The traditional mathematical techniques used in previous studies to model the wear of brake friction materials have not been able to understand the complex synergy arising from friction material properties.

Actually, machine learning algorithms address challenges with high accuracy and minimal computational costs, providing a distinct advantage over finite-element simulations. In contrast, finite-element simulations incur substantial computational expenses and time requirements [6]. The most effective method for forecasting is machine learning, in which the response is regulated by multiple parameters. Linear regression is a commonly employed machine learning technique for modeling and predicting linear correlations between explanatory factors (features) and target variable [7]. Multiple linear regression (MLR) analysis is a technique that allows the introduction of additional elements into the study to evaluate the influence of each one on the predicted target. This approach is valuable for assessing the impact of multiple concurrent factors on a single dependent variable [8] [9]. Multiple linear regression, according to Mata [10], is an approach used to represent the linear connection between dependent variables and one or more independent variables.

A multi-target linear regression approach, used to predict coefficient values, employs two different algorithms, namely Random Forest and Gradient Boosting

Regression. The random forest (RF) method, functioning as an ensemble of learning approach, employs a decision tree classifier for integrated decision-making [11]. In comparison to other machine learning methods, this approach stands out for its low computational requirements and high precision. It also exhibits robustness in handling missing and unbalanced data [12]. Because of its excellent approximation capacity, machine learning has been frequently applied for the prediction of tribological performance [13]. Other research, [14] used artificial neural networks (ANN) to forecast the wear of three different brake pad materials employing variables including sliding speed, normal load and temperature. Gyurova et al. [15] demonstrated the possibility of using machine learning algorithm to learn the results of wear tests of twenty nine different brake pad materials and correctly reproduces the tribological performance of friction material. Recently, Ikpambese demonstrates that machine learning model surpasses the Artificial Neural Network (ANN) model by achieving lower error values by the prediction of the wear of six different brake pad materials [16]. In addition to the remarkable predictive capabilities exhibited by machine learning models, it has been a recent emergence of explanation methods (SHAP) designed to facilitate the interpretation of intricate learning models. Enhanced comprehension holds particular significance by offering insights into the influence of different inputs on predictive outcomes. A substantial body of literature has concentrated on features related to performance, seeking to ascertain the degree to which each property contributes to the prediction results [17]. This alignment with the objectives of materials science establishes the groundwork for structure-properties relationships. This inclination towards elucidation implies an advantageous fusion of methods and applications. Notably, feature selection techniques have been employed to refine predictive models in materials science [18]. These comprehensive approaches, prove valuable, for a broad analysis of datasets, in identifying the physiochemical characteristics, correlated with functional properties [19].

Previous cited research introduced learning algorithms derived from experimental tests on specific brake linings with well-defined formulations. While these algorithms demonstrated good predictive performance within the scope of the tested formulations, they exhibit significant limitations when applied to other types of brake linings. The primary issue is that these models are heavily reliant on the specific formulations of the brake linings used during training. As a result, their applicability is restricted to those specific materials and does not generalize well to brake linings with different compositions or properties. This limitation severely impacts the versatility and practical utility of these algorithms in real-world applications, where a wide variety of brake lining materials are used. Additionally, there is a notable gap in research applying the SHAP (SHapley Additive exPlanations) method to enhance the interpretation of predicted friction and wear results in brake lining materials. Without such interpretability, understanding the underlying factors influencing the predictions remains challenging, further complicating the application and trust in these models across diverse material types.

In this article, a multi-target linear regression approach based on Random Forest and Gradient Boosting Regression algorithms was used to predict the friction coefficient and wear rate values of brake lining materials. We also investigate the application of explainable artificial intelligence (XAI) model through a systematic approach using Shapely Additive explanations (SHAP) to discuss materials properties and braking parameters that more influence brake pad tribological performance.

2 Materials

Brake lining materials play a crucial role in the braking system of vehicles. The formulations of brake friction materials are intricate, designed to meet multiple performance criteria, including maintaining a stable friction coefficient and minimizing wear. In fact, brake lining materials are composite materials whose formulation includes several ingredients. Through the combination of these elements, material acquires properties that allow it to better respond to brake solicitation [3]. In this sense, several researchers have focused on modifying the formulation in order to acquire the desired properties of the materials [20]. In addition, nowadays, given the harmful impact of particles emitted by brake systems and their negative impact on the environment and human health, several researchers have focused on the use of natural elements in the formulation of brake materials. As a result, today, we find in the literature several formulations of brake linings and each one has its own physical and mechanical properties [21], which subsequently affect its tribological performance. The experimental study of the tribological behavior proves that a third body manifests at the disc/ pad interface that makes the understanding of phenomena more difficult and delicate. Obviously, the third body characteristics are directly related to materials properties and brake action [22].

In fact, mechanical and physical properties especially, compression modulus and density, have a significant impact on these aspects [23]. Compression modulus of the brake lining material is essential as it determines its ability to adapt to pressure variations during braking, thereby affecting the stability of the friction coefficient. Adequate compression modulus helps maintain effective interaction with the brake disc, ensuring consistent response during braking cycles. On the other hand, material density is also a critical parameter, influencing wear resistance. Appropriate density promotes the uniform stress distribution and optimal heat dissipation, minimizing excessive wear. In fact, achieving a balance between compression and density in brake lining materials is essential to ensure effective braking performance, with a stable friction coefficient and controlled wear.

Recently, numerical simulation of the braking system has attracted researchers, given the independence of laboratory tests and associated problems. In this field, we have started several simulations using finite element calculation software

MATLAB Software [24] and ANSYS Software [25] [26]. Nevertheless, given the significant calculation time as well as the unknown phenomena that occurred at the interface and their evolving characteristics over time, attention is focused on this research study, on the prediction of tribological performances of brake lining materials through machine learning method.

3 Proposed Explainable Multi-Target Regression Algorithm

3.1 Data Description

The standard machine learning process encompasses severalphases:

(1) Data collection

(2) Data preprocessing

(3) Model description

(4) Explanation method (Figure 1)

Then, the explanation provided by the model were examined by comparing them with the results of the actual functioning of the brake system. The aim was to assess the accuracy and alignment between the model's predictions to the real values referred to previous scientific results. In the following section, we will provide a detailed account of the steps that we implemented in our work. It will offer a comprehensive understanding of procedures and methodologies applied in our research.

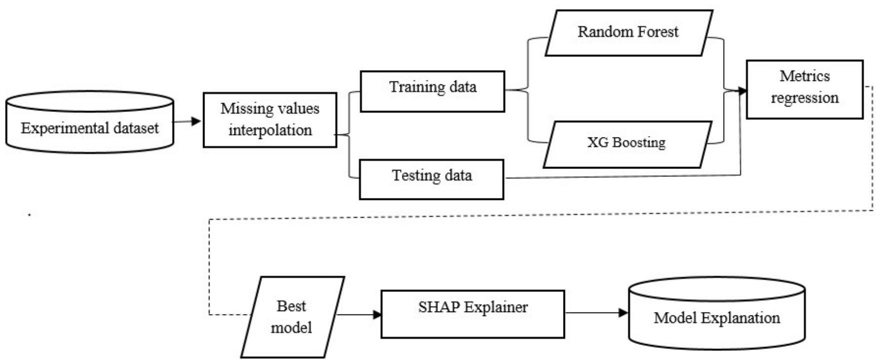


Figure 1
Proposed framework

3.2 Data Collection

The collected data encompasses information related to the characteristics of friction materials, as well as the specific conditions and variables manipulated during experimental procedures. The inclusion of both material properties and experimental parameters suggests a comprehensive approach to gathering information, aiming to capture a broad spectrum of factors that influence friction phenomena. Material properties, such as density and compression modulus of brake linings, alongside key experimental parameters like sliding velocity and pressure, serve as crucial input data for artificial intelligence models aiming to predict frictional performance. Table 1 presents different features and target values used in our model and their attributes. The density of brake material influences wear resistance, and the compression modulus indicates the material's ability to withstand braking pressures. In the other hand, sliding velocity helps to capture the impact of dynamic interactions on heat generation, wear rates, friction stability and pressure influences deformation and thermal response. Since their high impact on tribological performance, we have chosen to integrate compressive modulus, density, load and sliding speed into a machine learning model for the improvement of brake system performance.

Table 1

Dataset attributes detailed information

Attribute	Description	Type
comp	compression modulus	Input
density	density	Input
speed	sliding speed	Input
Load	Normal Load	Input
Friction	Friction coefficient	output
Wear	wear rate	output

Data collection is based on several researcher works. The advantage of gathering data from results of multiple researchers lies in the ability to cover a broad range of brake lining materials. By aggregating data from multiple scientific works, the representative of the entire spectrum of brake lining materials is enhanced, providing a more complete understanding of the relationships between material properties and frictional performance. This approach also takes account for significant variations in experimental conditions and material characteristics, thereby reinforcing the robustness and generality of predictive models developed from these combined datasets. Our research utilizes systematic mapping as a method for processing extracting materials properties and performance from previous research works, a technique developed and applied by professionals [27]. Given the extensive volume of research articles in the field of friction and wear performances of brake pad materials, the selection of a reliable database becomes crucial. In this regard, scholars in the broader domain of utilizing Machine Learning

(ML) for concrete characterization commonly advocate for using Scopus bibliometrics records [28].

To construct the annotated database for training, we first conducted a comprehensive literature review to identify relevant scientific papers detailing friction and wear measurements. We systematically extracted data points, ensuring key experimental details such as material properties (compression modulus and density), testing conditions (load, speed), and measurement outcomes (friction coefficient and wear volume). The extracted data underwent preprocessing, including error removal, consistency checks, and normalization to ensure comparability across different scales of measurement (for some cases, we calculated load from pressure or converted units of sliding speed from rev/min to m/s). Then, the dataset was organized into structured file formats (CSV). The dataset comprises a total 171 instances. All instances used in the model comprise material properties and experimental parameters as input and tribological performance as output. In Table 2, we present the number of instances used for each research work. It provides an illustrative representation of tribological test parameters and properties of brake lining materials serving as features, with the corresponding tribological performance of brake pad material serving as the target value for each scientific work. Approximately 80% of the total data was used for training purposes, while the remaining 20% was reserved for assessing the quality of algorithms predictions [29].

Table 2
An excerpt of the collected dataset

Number of instances	**Comp (MPa)**	**Density (g/cm3)**	**Load (N)**	**Speed (m/s)**	**Friction**	**Wear (10^{-6} g/m)**	**Ref**
12	87	0.9	20	5.02	0.29	6.1	[30]
-9	110	1.89	100	6	0.35	3.8	[31]
8	110	1.89	30	4	0.32	3	[32]
3	55	2.22	30	1.8	0.51	10.55	[33]
6	65.11	3.704	41	22	0.2924	3.875	[34]
8	110	NaN	NaN	NaN	0.35	3.8	[35]
5	110	1.89	NaN	NaN	0.35	3.8	[36]
7	112	1.429	20	21	0.43	3.48	[37]
7	110	2.06	NaN	NaN	0.35	3.8	[38]
4	NaN	2.247	100	1.33	0.36	1.0749	[39]
22	211	2.4	300	6.7	0.57	2.808	[40]
5	105	1.43	20	5.02	NaN	4.2	[41]
8	189	NaN	NaN	NaN	0.35	3.8	[42]
27	NaN	1.566	9.81	1.41	0.901	3	[43]
4	NaN	2.17	101	6	0.3	4.2	[44]
27	101	2.592	8	0.1	0.88	1.08	[45]
6	110	1.89	NaN	NaN	NaN	3.8	[46]
6	87.9	1	20	5.02	NaN	6.2	[47]

3.2 Data Preprocessing

Data preprocessing is crucial to ensure the success of machine learning models, and addressing missing data is a pivotal component of this process. Missing data poses significant challenges to the performance and robustness of predictive models. It introduces the risk of bias, as models may learn from incomplete and unrepresentative samples, leading to skewed inferences. The reduction in the effective sample size can diminish the model's statistical power and the ability to generalize accurately to new data. Unbalanced representations, instability, and biased inferences are among the potential consequences. Additionally, the impact on feature importance and the overall effectiveness of generalization can be compromised if missing data is not appropriately addressed. Handling missing data is a critical aspect of preprocessing to ensure the reliability and fairness of machine learning models. Missing data are presented in Table 2, as NaN associated to specific material properties, commonly linked to the compression modulus, as researchers might emphasize hardness in their work and may not encompass physical properties like density. In certain cases, scientists may concentrate solely on friction coefficient or wear rate, resulting in the absence of some data related to material characteristics and performances. When dealing with missing data, which are independent of the observed and unobserved data, we estimate the missing values using means or average values. This approach is known as missing completely at random (MCAR) and it involves replacing the missing data points with the mean (average) of the observed values [48].

3.4 Model Description

After completing the data preprocessing phase's, two frequently used machine learning algorithms, namely, Gradient Boosting (XG) and random forest (RF) were used to predict target values (friction coefficient and wear rate). RF is an ensemble of learning method that operates by constructing a multitude of decision trees during training. In classification tasks, the class selected by the majority of trees determines the final output. For regression tasks, the output is the average of the output values from different trees [49]. This method is computationally efficient and exhibits high speed and accuracy [50]. Gradient Boosting, another machine learning algorithm employed for both classification and regression tasks, constructs a robust predictive model by amalgamating the predictions from multiple weak learners that focus on the mistakes made by the existing ensemble, often in the form of decision trees. Gradient Boosting Regression can be extended to predict multiple variables simultaneously, often referred to as a multivariate regression task. Most popular implementations of Gradient Boosting, such as Xgboost, naturally support multivariate regression. The approach's strength lies in its capability to concurrently capture intricate relationships between input variables and multiple target variables.

3.5 Explanation Method

Numerous methods have been proposed to elucidate predictions made by machine learning models. SHAP is considered as the most dominant approaches. The SHAP method is a game theory-based explainable artificial intelligence approach designed to calculate shapely values, assessing the impact of each feature on the prediction. [51]. In computing shapely values, SHAP selectively includes some feature values while excluding others, with the primary aim of discerning each feature's contribution to the prediction.

Some researchers [51] developed a python package capable of computing SHAP values for various technologies, such as Xgboost, and tree models. This implementation has gained widespread adoption among researchers who rely on SHAP for interpreting various models [52]. Additionally, Lundberg et al. [53] proposed an extension of SHAP known as the SHAP tree explainer. This extension suggests that the precise evaluation of SHAP values can be achieved in polynomial time exclusively for tree-based models, including Random Forest (RF). It is valuable for explaining both regression and classification models, as well as complex algorithms like Random Forest (RF) and Gradient Boosting (XG).

Our investigation deals with the explanation of a black-box model in friction and wear prediction. We focused on the understanding of rules learned by the machine learning model, particularly the importance and influence of predictor variables (compression modulus, density, sliding speed and load) on target values, namely wear rate and friction coefficient.

3.6 Regression Metrics

The performance of the model is assessed using the following regression metrics:

MSE, which stands for Mean Squared Error, assesses the average squared disparity between predicted and actual values. It provides a quantitative measure of how well a regression model performs in terms of accuracy. The formula for MSE_j is for the jth target as follows Eq 1

$$MSE_j = \sum_{i=1}^{n} (yij(pred) - yij(real))^2 \tag{1}$$

MAE, Mean Absolute Error, evaluates the average absolute discrepancy between predicted and actual values. It provides a straight forward and easy to interpret measure of the model'sperformance in terms of accuracy. The formula for MAE_j for the jth target is as follows Eq 2

$$MAE_j = \sum_{i=1}^{n} |yij(pred) - yij(real)| \tag{2}$$

R-squared (R^2) represents the coefficient of determination.It assesses the proportion of variance in the dependent variable that can be explained by the independent variables in a regression model. The formula for $R_j{}^2$ for the *j*th target is as follows Eq 4:

$$R_j{}^2 = \sum_{i=1}^{n} \frac{(yij(pred) - yij(real))^2}{|yij(pred) - yij(real)|} \tag{4}$$

With:

n is the Number of Instances

yij(pred) is the predicted value of *j*th target variable for the i th instance

yij(real) is the real value of the *j*th target variable for the i th instance

After calculating regression metrics values for friction coefficient and wear, we employed an approach that combines these measurements into a single performance metric by averaging them.

4 Results

Figure 2 displays a comparison of regression metric results for two algorithms, specifically focusing on MSE (Figure 2a), MAE (Figure 2b), and R-squared values (Figure 2c). In Figure 2a, MSE for XGBoost (XG) is reported as 0.857, while for Random Forest (RF), MSE is approximately 1.2. This indicates a substantial difference, with XG achieving a significantly lower MSE compared to RF. A lower MSE suggests that XGBoost has a better capability to minimize the squared differences between predicted and actual values, thus providing more accurate predictions.

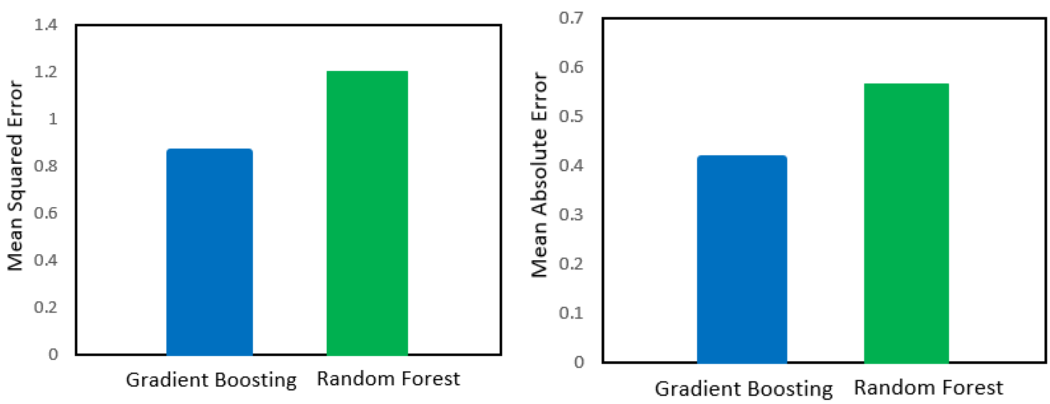
Furthermore, when examining the MAE, the XG algorithm demonstrates a value of 0.4138, which is notably lower than the MAE of 0.567 for RF (Figure 2b). A lower MAE indicates that XGBoost produces predictions that are closer to the actual values on average, suggesting that it makes smaller errors in prediction compared to RF. This is crucial in the friction and wear prediction task, where precise predictions can significantly impact the outcomes.

In terms of R-squared values, XG exhibits a higher value of 0.756, whereas RF shows a value of 0.59 (Figure 2c). A higher R-squared value indicates that a greater percentage of the variance in the dependent variable is accounted for by the independent variables in the model, highlighting XGBoost's superior explanatory power. This suggests that XGBoost provides a better fit to the underlying data structure, capturing the relationships between features more effectively than RF.

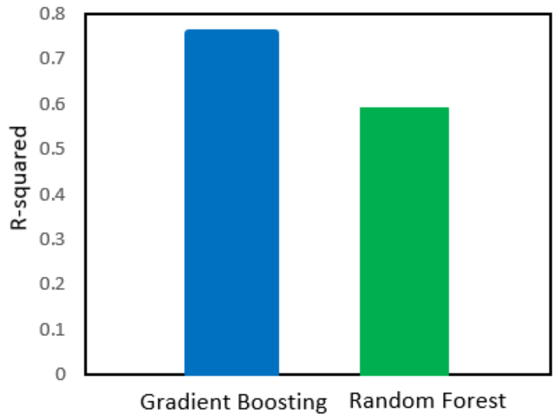
Additionally, the interpretability of these models was examined using SHAP (SHapley Additive exPlanations) values. SHAP values provide a breakdown of each feature's contribution to the model's predictions, offering insights into the decision-making process of the model. The SHAP analysis revealed that certain features, such as the hardness of materials and sliding speed, had a more significant influence on the predictions in the XGBoost model than in the RF model. This detailed understanding of feature importance is essential for interpreting the model's behavior and for making informed decisions based on its predictions.

Furthermore, XGBoost's ability to handle missing data and model intricate interactions between variables, as evidenced by the SHAP analysis, underscores its robustness and flexibility in dealing with complex datasets. This aligns with the findings of another researcher referenced as [54], which supports the idea that Extreme Gradient Boosting can be highly effective in predicting time-series tabular data. The consistency between our model results and those of the referenced study adds credibility to the claim that XGBoost demonstrates superior performance in predictive tasks.

Overall, the results presented in Figure 2 consistently indicate that, across multiple metrics—MSE, MAE, and R-squared—XGBoost outperforms Random Forest in the specified prediction scenario. The detailed analysis provided by SHAP values further reinforces this conclusion, showcasing XGBoost's superior feature interpretation capabilities and robustness in handling complex interactions. This comprehensive evaluation emphasizes the potential of XGBoost as a robust predictive model, particularly in tasks involving intricate data patterns.



a) b)



c)

Figure 2
Comparison results for XG (blue) and RF (green) a) MSE b) MAE c) R^2

The information from Figures 3 and 4 indicates that the comparison between the predicted values of wear rate and friction coefficient generated by the XG model and real values reveals a noteworthy agreement. In fact, both the test set error and the validation set error share similar characteristics. Notably, there is no substantial evidence of over fitting observed across iterations. The observed closeness between predicted, and experimental values proves the ability of the XG model in delivering precise predictions for friction coefficient and wear rate in the field of brake pad materials. Overall, these findings underscore the model's stability, generalization capacity, and accuracy in predicting target variables, thereby strengthening confidence in its overall performance and reliability.

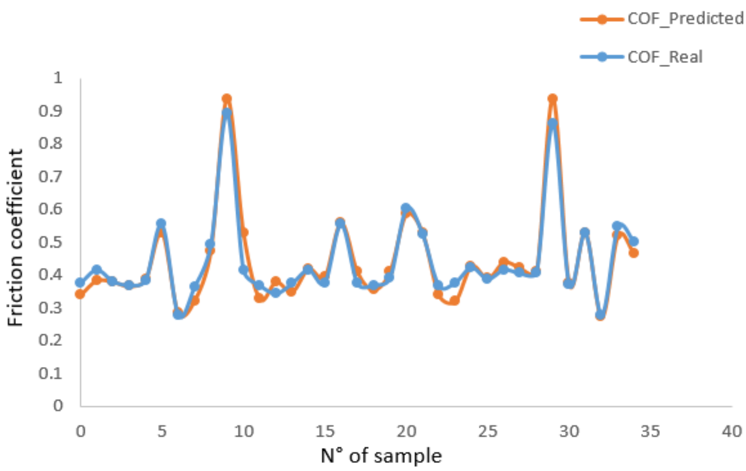


Figure 3
Comparison between predicted friction coefficient and real values

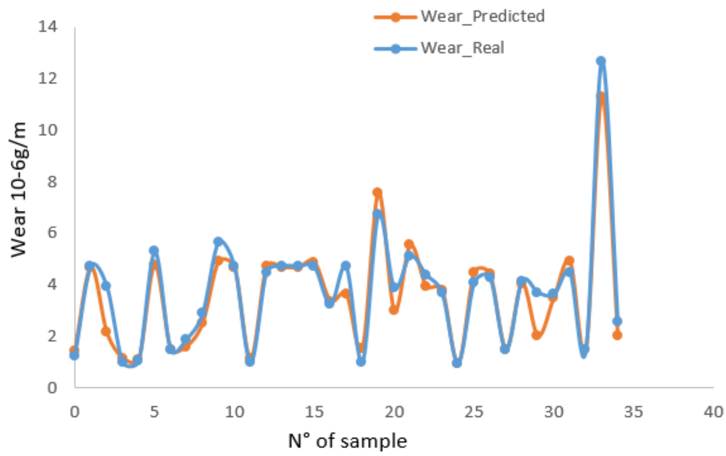


Figure 4
Comparison between predicted wear rate and values

The SHAP summary plot is shown in Figure 5. On the y-axis, feature names are presented in descending order of significance, while the SHAP values for each input predictor are depicted on the x-axis. Results shows that sliding speed is the most influential parameters. The emphasis on sliding speed underscores its pivotal role in tribological performance prediction, signaling its dominance among the factors considered in the model. This recognition is essential for tailoring strategies to optimize tribological outcomes in various applications. This finding is also reported by other research work, proving that, sliding speed has the most significant influence on tribological performance prediction [55].

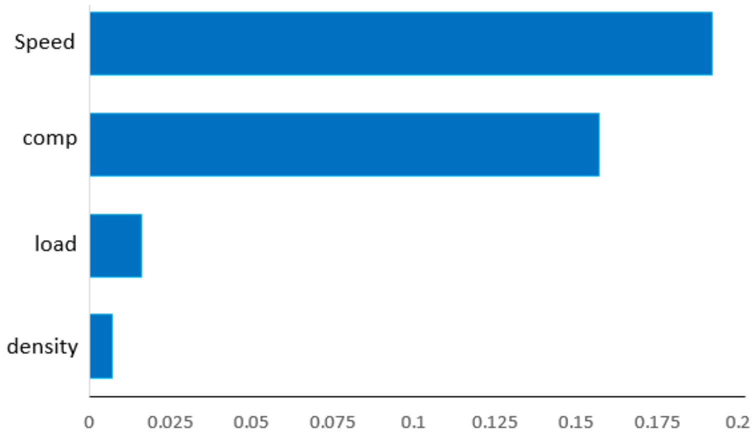


Figure 5
Mean SHAP: average impact on friction and wear magnitude

In Figure 6, red points indicate higher feature values, whereas blue points signify lower values. The SHAP summary plots showed the distribution of SHAP values for each feature, highlighting the features with the highest average impact on the models. This visualization offers a comprehensive overview of the XG model by emphasizing the importance of each feature and illustrating its impact on the model's outputs.

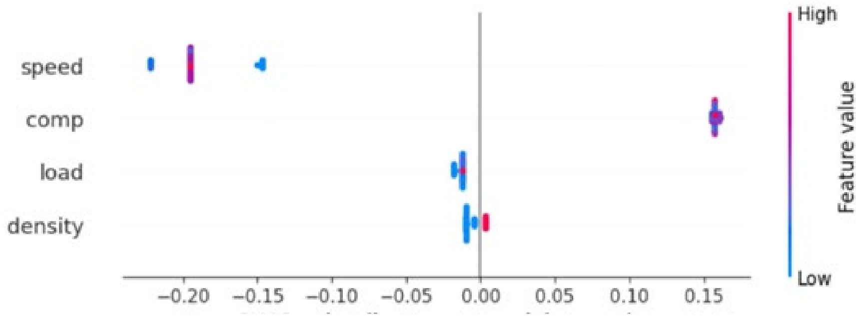


Figure 6
SHAP Summary Plot of XG Model

According to Figure 6, features linked to properties of materials such, compression modulus and density have a positive overall impact on the model's output, unlike those of test parameters such as speed and load that have a negative impact. In other words, high values of the first two features (compression and density) are associated with low wear rate, whereas high values of the last two (sliding speed and load) are associated with reduced friction coefficient and high wear rate. This explanation is critical because it demonstrates that the model is correctly learning the dynamics of friction coefficient. For example, it learned that brake materials with high compression modulus, impact a reduced wear rate. In fact, since its stiffness, on a few quantities of third body formed by the friction of brake pad materials. A stable friction film was generated on the surface of the brake pad, which provides excellent friction stability with reduced wear rate [44].

Furthermore, the high value of test parameters; sliding speed and load, knowing as parameters linked to severe brake solicitation studied by Sellami et al, induces low level of friction coefficient. She proves that at this level of solicitation, friction coefficient is reduced and, the phenomena is called "fade" associated with pronounced wear rate [56]. Recently, Kenneth M. Jensen et al. [57] proved that brake pad wear is proportional to sliding speed and load, and inversely proportional to the hardness of the brake pad. These findings confirm the SHAP results.

Conclusions

Developing models, related to the tribological performance of brake friction materials is crucial, due to the increasing complexity of requirements imposed on these materials. Brake pads, in particular, undergo intricate tribological interactions, given the diverse mechanical properties of their ingredients and the varying conditions during the brake operation. An approach has been proposed to model the influence of pertinent factors linked to the friction material properties: compression modulus and density and to brake solicitation: sliding speed and applied load. The utilization of a computer-based model through multi target regression employing Random Forest and extreme Gradient Boosting has been suggested to address the challenge of predicting friction coefficient and wear rate for brake friction materials. A total of 171 instances were investigated, referring to research done by scientific specializing in brake lining materials. The XG algorithm demonstrated the best results in predicting tribological behavior. The developed XG model successfully predicted wear and friction across various types of friction materials. The multi-target regression model was trained to understand how the wear and friction performance of friction materials are influenced by material's properties and braking solicitations.

The analysis indicated that the sliding speed was the parameter exerting the most significant influence and concurrently, the compression modulus that emerges as the most pivotal of the material's properties. The SHAP models provided insights into the dynamics, illustrating that material properties exert a positive impact on tribological performance. In contrast, parameters related to brake solicitation were revealed to have a detrimental effect on material performance.

The correlation between predicted and real friction and wear rate results, along with the SHAP findings, align well with the analytical and experimental results from other researchers. This underscores the potential of the proposed XG algorithm for forecasting tribological performance of brake lining materials in the future.

By combining the predictive power of XGBoost with the interpretative insights from SHAP, we achieved a comprehensive understanding of the factors influencing friction and wear. This approach not only improved prediction accuracy but also provided valuable insights into the underlying mechanisms, facilitating better decision-making and optimization in tribological applications.

To improve the precision of the multi-target regression model for predicting the tribological performance of brake pad materials, one recommended approach is to increase the quantity of output data, associated with wear and friction coefficient measurements in diverse wear tests.

Expanding the dataset in this manner allows the model to capture a more comprehensive range of scenarios and variations in material behavior under different conditions. This not only enhances the model's ability to generalize, but also provides a more robust foundation for understanding the nuanced relationships between material properties and tribological outcomes. Additionally, the increased data diversity aids in refining the model's predictive capabilities, ultimately contributing to a more accurate representation of the complex dynamics involved in brake pad performance.

Acknowledgements

This work was supported by the Tunisian Ministry of Higher Education and Scientific Research for their continuous support at the Electro-Mechanical Systems Laboratory of Sfax and Higher Institute of Transport and Logistics of Sousse.

References

[1] P. Zhang, L. Zhang, D. Wei, P. Wu, J. Cao, C. Shijia, and X. Qu, ‘‘A high-performance copper-based brake pad for high-speed railway trains and its surface substance evolution and wear mechanism at high temperature,’’ Wear, Vol. 444-445, p. 203182, 2020

[2] A. Sellami, M. Kchaou, R. Kus, J. Fajoui, R. Elleuch, and F. Jacquemin, ‘‘Impact of brass contents on thermal, friction and wear properties of brake linings composites,’’ Mechanics & Industry, Vol. 19, p. 105, 2018

[3] A. Sellami, N. Hentati, M. Kchaou, M. A. Chowdhury, and R. Elleuch, ‘‘Effect of size and shape of copper alloys particles on the mechanical and tribological behavior of friction materials,’’ Mechanics & Industry, 2020

[4] X. Yan, D. Lin, B. Chen, Y. Ding, N. She, and F. Zhang, ‘‘Study on the influence of three-body particles on the dynamic performance of braking system,’’ Tribology International, Vol. 189, p. 109013, 2023

[5] Z. Xiang, J. Mo, H. Ouyang, F. Massi, B. Tang, and Z. Zhou, ‘‘Contact behaviour and vibrational response of a high-speed train brake friction block,’’ Tribology International, Vol. 152, p. 106540, 2020

[6] P. Grzes and M. Kuciej, ‘‘Coupled thermomechanical fe model of a railway disc brake for friction material wear calculations,’’ Wear, Vol. 530-531, p. 205049, 2023

[7] S. Wang, S. Khatir, and M. Abdel Wahab, ‘‘Proper orthogonal decomposition for the prediction of fretting wear characteristics,’’ Tribology International, Vol. 152, p. 106545, 2020

[8] S. Goyal and G. K. Goyal, ‘‘Application of artificial neural engineering and regression models for forecasting shelf life of instant coffee drink,’’ 2011

[9] A. Livinus, ‘‘Predictive models for oil in place for oil rim reservoirs in the niger delta using machine learning approach,’’ Petroleum Petrochemical Engineering Journal, Vol. 7, pp. 1-16, 07 2023

[10] J. Mata, ‘‘Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models,’’ Engineering Structures, Vol. 33, No. 3, pp. 903-910, 2011

[11] A. Cutler, D. Cutler, and J. Stevens, Random Forests, Vol. 45, pp. 157-176, 01 2011

[12] M. Córdoba and M. Balzarini, ''A random forest-based algorithm for data-intensive spatial interpolation in crop yield mapping,'' Computers and Electronics in Agriculture, Vol. 184, p. 106094, 05 2021

[13] U. M. R. Paturi, S. T. Palakurthy, and N. S. Reddy, ''The role of machine learning in tribology: A systematic review,'' Archives of Computational Methods in Engineering, Vol. 30, pp. 1-53, 11 2022

[14] D. Aleksendrić, ''Neural network prediction of brake friction materials wear,'' Wear, Vol. 268, No. 1, pp. 117-125, 2010

[15] L. A. Gyurova and K. Friedrich, ''Artificial neural networks for predicting sliding friction and wear properties of polyphenylene sulfide composites,'' Tribology International, Vol. 44, No. 5, pp. 603-609, 2011, Special Issue: ECOTRIB 2009

[16] K. Ikpambese, E. Lawrence, and L. Enyejo, ''Tribology in industry comparative analysis of multiple linear regression and artificial neural network for predicting friction and wear of automotive brake pads produced from palm kernel shell keywords: Multiple linear regression artificial neural network wear rate coefficient of friction inertia dynamometer,'' pp. 565-573, 01 2018

[17] A. Altmann, L. Tolosi, O. Sander, and T. Lengauer, ''Permutation importance: A corrected feature importance measure,'' Bioinformatics (Oxford, England), Vol. 26, pp. 1340-7, 04 2010

[18] Y. Xu, L. Jiang, and X. Qi, ''Machine learning in thermoelectric materials identification: Feature selection and analysis,'' Computational Materials Science, Vol. 197, p. 110625, 2021

[19] T. Liu and A. S. Barnard, ''The emergent role of explainable artificial intelligence in the materials sciences,'' Cell Reports Physical Science, Vol. 4, No. 10, p. 101630, 2023

[20] M. Baklouti Kharrat, R. Elleuch, C. Anne Lise, D. Najjar, and Y. Desplanques, ''Relationships between the heterogeneous microstructure, the mechanical properties and the braking behaviour of an organic brake lining material,'' Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, Vol. 227, pp. 549-560, 04 2013

[21] E. R. Sellami Amira, ''Green composite friction materials: A review of a new generation of eco-friendly brake materials for sustainability,'' Environmental Engineering Research, Vol. 29, No. 3, pp. 230226-0, 2024

[22] A. Sellami, N. Hentati, M. Kchaou, R. Kus, R. Elleuch, and M. A. Chowdhury, ''Wear mechanisms of different engineering systems under higher solicitations: Overview and case studies,'' Engineering Failure Analysis, Vol. 94, pp. 165-181, 2018

[23] M. Yigrem, O. Fatoba, and S. Tensay, ‘‘Tribological and mechanical properties of banana peel hybrid composite for brake-pad application,’’ Materials Today: Proceedings, Vol. 62, pp. 2829-2838, 2022, International Conference on Materials, Processing Characterization (13th ICMPC)

[24] A. Sellami, M. Kchaou, R. Elleuch, and Y. Desplanques, ‘‘Thermal analysis of pad-on-disc contact under tribological solicitations: a coupled numerical-experimental approach to identify surface temperatures and flow partition coefficient,’’ Heat and Mass Transfer, Vol. 52, pp. 1923-1934, Sept. 2016

[25] S. Amira, Z. Kawther, and E. Riadh, ‘‘Towards the development of a numerical model for the simulation of thermal behavior of disc brake systems,’’ in Advances in Acoustics and Vibration IV (A. Akrout, M. Abdennadher, N. Feki, M. S. Abbes, F. Chaari, and M. Haddar, eds.), (Cham), pp. 433-439, Springer International Publishing, 2023

[26] S. Amira, G. M. Hedi, and E. Riadh, ‘‘Enhanced disc brake thermo-kinetics for better wear test reproducibility,’’ in Advances in Acoustics and Vibration IV (A. Akrout, M. Abdennadher, N. Feki, M. S. Abbes, F. Chaari, and M. Haddar, eds.), (Cham), pp. 425-432, Springer International Publishing, 2023

[27] Y. Pan, L. Deng, S. Li, J. Wang, and F. Zhang, ‘‘A study using a combined method of scientometric and manual analysis to review the present research of plant fibres reinforced concrete,’’ Construction and Building Materials, Vol. 341, p. 127551, 2022

[28] E. M. Lasda Bergman, ‘‘Finding citations to social work literature: The relative benefits of using web of science, scopus, or google scholar,’’ The Journal of Academic Librarianship, Vol. 38, No. 6, pp. 370-379, 2012

[29] E. B. Abdallah, R. Grati, and K. Boukadi, ‘‘Towards an explainable irrigation scheduling approach by predicting soil moisture and evapotranspiration via multi-target regression,’’ J. Ambient Intell. Smart Environ., Vol. 15, pp. 89-110, 2023

[30] U. Idris, V. Aigbodion, I. Abubakar, and C. Nwoye, ‘‘Eco-friendly asbestos free brake-pad: Using banana peels,’’ Journal of King Saud University - Engineering Sciences, Vol. 27, No. 2, pp. 185-192, 2015

[31] R. Rao and B. Gudla, ‘‘A review paper on alternate materials for asbestos brake pads and its characterization,’’ Vol. 2, pp. 556-562, 11 2022

[32] O. Adeyemi, N. Ademoh, and T. Boye, ‘‘Development of asbestos-free automotive brake pad using ternary agro-waste fillers,’’ Journal of Multidisciplinary Engineering Science and Technology (JMEST), Vol. 3, pp. 5307-5323, 07 2016

[33] A. P. N. Elzayady, ‘‘Manufacturing and development of low cost asbestos-free brake pad composite material,’’ 08 2021

[34] N. EL-Tayeb and K. Liew, ‘‘On the dry and wet sliding performance of potentially new frictional brake pad materials for automotive industry,’’ Wear, Vol. 266, No. 1, pp. 275-287, 2009

[35] A. P. Irawan, D. F. Fitriyana, C. Tezara, J. P. Siregar, D. Laksmidewi, G. D. Baskara, M. Z. Abdullah, R. Junid, A. E. Hadi, M. H. M. Hamdan, andN. Najid, ‘‘Overview of the important factors influencing the performance of eco-friendly brake pads,’’ Polymers, Vol. 14, p. 1180, March 2022

[36] U. Idris, V. Aigbodion, I. Abubakar, and C. Nwoye, ‘‘Eco-friendly asbestos free brake-pad: Using banana peels,’’ Journal of King Saud University - Engineering Sciences, Vol. 27, No. 2, pp. 185-192, 2015

[37] A. Madeswaran, B. Natarajasundaram, and B. Ramamoorthy, ‘‘Reformation of eco-friendly automotive brake pad by using natural fibre composites,’’ 2016

[38] G. P. Williams, ‘‘Reference module in materials science and materials engineering,’’ 1970

[39] F. Yusubov, ‘‘Influence of the particle size on the physical, mechanical and tribological properties of composites for brake pads,’’ Vestnik Mashinostroeniya, pp. 36-40, 05 2021

[40] B. Öztürk, F. Arslan, and S. Öztürk, ‘‘Effects of different kinds of fibers on mechanical and tribological properties of brake friction materials,’’ Tribology Transactions, Vol. 56, 07 2013

[41] A. V.S, U. Akadike, S. Hassan, F. Asuke, and J. Agunsoye, ‘‘Development of asbestos - free brake pad using bagasse,’’ Tribology in Industry, Vol. 32, pp. 12-18, 03 2010

[42] O. Adeyemi, N. Ademoh, and M. Okwu, ‘‘Development and assessment of composite brake pad using pulverized cocoa beans shells filler,’’ International Journal of Materials Science and Applications, Vol. 5, p. 66, 01 2016

[43] J. Sachin, S. H. S. Sanjay, D. G. B. Digvijay, D. Y. Patil, and E. Society, ‘‘Characterization of newly developed asbestos and copper-free low metallic friction material for automotive brake pad application,’’

[44] M. Kchaou, A. Sellami, R. Elleuch, and H. Singh, ‘‘Friction characteristics of a brake friction material under different braking conditions,’’ Materials & Design, Vol. 52, pp. 533-540, 2013

[45] G. Friday, A. Bello, A. Tokan, and B. A, ‘‘Physico-mechanical properties of basalt-based brake pad as alternative to ceramics brake pad,’’ Saudi Journal of Engineering and Technology, Vol. 7, pp. 16-33, 01 2022

[46] V. Haruna, A. S. Abdulrahman, M. S. Abolarin, and R. A. Muriana, ‘‘Evaluation of sheanut shell-reinforced automotive brake pad,’’ Arid Zone Journal of Engineering, Technology and Environment, Vol. 15, pp. 510-518, 2019

[47] A. Ohueyi Ahmed, A. Umar, B. U. Usman, E. Aliyu, S. Omilabu, and S. Khalifa, ''Development of asbestos free-brake pad using solid waste,'' Journal of Engineering, 06 2018

[48] P. C. Austin, I. R. White, D. S. Lee, and S. van Buuren, ''Missing data in clinical research: A tutorial on multiple imputation,'' The Canadian Journal of Cardiology, Vol. 37, pp. 1322-1331, 2020

[49] K. Satpathi, V. Balijepalli, and A. Ukil, ''Modeling and real-time scheduling of dc platform supply vessel for fuel efficient operation,'' IEEE Transactions on Transportation Electrification, Vol. 3, pp. 762-778, 08 2017

[50] M. Córdoba and M. Balzarini, ''A random forest-based algorithm for data-intensive spatial interpolation in crop yield mapping,'' Computers and Electronics in Agriculture, Vol. 184, p. 106094, 05 2021

[51] S. M. Lundberg and S.-I. Lee, ''A unified approach to interpreting model predictions,'' in Neural Information Processing Systems, 2017

[52] S. Ergün, ''Explaining xgboost predictions with shap value: a comprehensive guide to interpreting decision tree-based models,'' New Trends in Computer Sciences, Vol. 1, pp. 19-31, 04 2023

[53] S. M. Lundberg, G. G. Erion, and S.-I. Lee, ''Consistent individualized feature attribution for tree ensembles,'' ArXiv, Vol. abs/1802.03888, 2018

[54] D. Chakraborty, H. Başağaoğlu, and J. Winterle, ''Interpretable vs. noninterpretable machine learning models for data-driven hydro-climatological process modeling,'' Expert Systems with Applications, Vol. 170, p. 114498, 2021

[55] A. Senatore, V. D'Agostino, R. Di Giuda, and V. Petrone, ''Experimental investigation and neural network prediction of brakes and clutch material frictional behaviour considering the sliding acceleration influence,'' Tribology International, Vol. 44, No. 10, pp. 1199-1207, 2011

[56] A. Sellami, M. Kchaou, R. Elleuch, A.-L. Cristol, and Y. Desplanques, ''Study of the interaction between microstructure, mechanical and tribo-performance of a commercial brake lining material,'' Materials Design, Vol. 59, pp. 84-93, 2014

[57] Kenneth M. Jensen, Ilmar F. Santos, Harry J. P. Corstens, "Estimation of brak pad wear and remaining useful life from fused sensor system, statistical data processing, and passenger car longitudinal dynamics", Wear, Vols. 538-539, 2024, https://doi.org/10.1016/j.wear.2023.205220