Engineering Journal
issn: 0125-8281

Article

Improving Aggregate Abrasion Resistance Prediction via Micro-Deval Test Using Ensemble Machine Learning Techniques

Alireza Roshan[1,a,*] **and Magdy Abdelrahman**[2,b]

1 Department of Civil, Architectural and Environmental Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA
2 Missouri Asphalt Pavement Association (MAPA) Endowed Professor, Department of Civil, Architectural and Environmental Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA
E-mail: [a,*]alireza.roshan@mst.edu (Corresponding author), [b]m.abdelrahman@mst.edu

Abstract. Aggregate is the most extracted material from the world's mines and widely used in civil and construction projects. The Micro-Deval abrasion test (MD) is one of the most important tests that provides characteristics of crushed aggregates that show their resistance against mechanical abrasive factors such as repeated impact loading. The impact of various factors on abrasive resistance properties of aggregates has led researchers to seek correlations, often focusing on limited data samples, leading to reduced accuracy. This study employs machine learning (ML) methods to predict MD abrasion values, considering diverse aggregate properties. Various ensemble ML methods were applied, revealing the exceptional performance of the stacking model, which achieved an R^2 score of 0.95 in predicting aggregate abrasion resistance. The feature importance analysis highlights the influence of factors such as Magnesium Sulfate Soundness (MSS), Water Absorption (ABS), and Los Angeles Abrasion (LAA) on aggregate abrasion values, suggesting that the use of multiple test methods could yield a more dependable assessment of aggregate durability.

Keywords: Aggregate abrasion resistance and durability, micro-Deval abrasion test, friction assessment, ensemble machine learning.

ENGINEERING JOURNAL Volume 28 Issue 3
Received 16 November 2023
Accepted 15 March 2024
Published 31 March 2024
Online at https://engj.org/
DOI:10.4186/ej.2024.28.3.15

DOI:10.4186/ej.2024.28.3.15

1. Introduction

Aggregates, a blend of fine and coarse particles including gravel, sand, crushed stone, slag, concrete waste, polymer materials, glass, and rubber fragments, are essential in various structures. Its primary use is found in the construction of concrete structures, the surfacing of roads, and the development of railway infrastructure, where it often makes up a significant part of the composition, representing approximately 70-80% of the material in concrete.

deployed in a specific grain size distribution, they directly impact characteristics like compressive strength, thermal expansion, and concrete density. In road construction, asphalt, or asphaltic concrete, stands as the most prevalent material. Aggregates, acting as the primary load-bearing elements, make up approximately 90-95% by weight and 75-85% by volume of asphalt compositions. Whether sourced from natural deposits, processed from broken materials, or synthetically manufactured, aggregates must consistently conform to rigorous requirements for road pavement, encompassing physical attributes such as granularity, durability, toughness, shape, texture, water absorption, and abrasion [1].

Insufficient friction between vehicle tires and road surfaces is a significant contributor to traffic accidents. To mitigate accidents and fatalities arising from poor friction, optimal skid resistance must be ensured under various weather conditions. The primary determinants of high pavement friction are the type and texture of aggregates. In practical terms, addressing this issue requires conducting laboratory testing and evaluating aggregate materials. Two widely used tests for ensuring the quality control of aggregates include the Micro-Deval abrasion test (MD) and the Los Angeles abrasion test (LAA) [2, 3]. The MD test, developed in France in the 1960s, evaluates mineral aggregates' abrasion resistance and durability by simulating abrasive actions with steel balls in the water. This test is highly valuable as it closely mimics the conditions experienced in the field, where aggregates are subject to abrasive forces in the presence of moisture. In reality, water can infiltrate the interface between aggregates and asphalt, potentially causing adhesion loss and asphalt concrete pavement failure. Therefore, it's crucial to consider moist conditions resulting from rain, snow, and ice melt during pavement service [4,5]. Despite the widespread use of the LAA test for evaluating aggregate durability, many researchers have expressed reservations about its suitability for assessing aggregates intended for use in asphalt and concrete pavements [6]. This discrepancy stems from variations in the behavior of different aggregate materials; those with strong crystals like granite may respond differently than materials with weaker crystals, such as slates, during the LAA test [7]. In response to the concerns mentioned, the American Association of State Highway and Transportation Officials (AASHTO), as part of the National Cooperative Highway Research Program (NCHRP-405) project, investigated aggregate behavior in asphalt pavements. The primary objectives were to analyze and compare outcomes from the LAA and MD tests and establish correlations with assessments like the magnesium sulfate soundness test [8]. Testing 16 aggregate types revealed a strong correlation between the MD test and the magnesium sulfate soundness test, providing valuable insights into the aggregate performance in asphalt pavements. In a study by M. Takarli, a strong link was found between aggregate abrasion resistance and mineral composition. Their findings indicated a significant correlation between mineral compositions and various aggregate properties, including density, water absorption, and wear resistance [9]. This research illuminated how an aggregate's mineral composition influences its physical characteristics and durability, providing valuable insights for material selection and pavement design. Another study reviewed the relationship between the geology of aggregates and their performance in LAA and MD tests. It was demonstrated that drawing final conclusions about mechanical performance based on a single textural factor is not adequate. Therefore, a comprehensive consideration of all relevant factors simultaneously is necessary [10]. The demand for natural aggregates has experienced a substantial surge due to the expansion of construction projects and the depletion of natural aggregate reserves. This has prompted researchers to explore alternatives, such as recycled aggregates, and seek ways to minimize destructive testing of granular materials. Consequently, computer-based methods have gained prominence in the realm of engineering sciences [11]. Recently, artificial intelligence methods have experienced substantial growth in addressing large-scale challenges within the engineering sciences. The rise in available data and the efficiency of modern computational processes have fueled the adoption of ML methods, especially in fields like geotechnical engineering and material science which have attracted considerable attention over the past decade [12].

Researchers have widely applied ML to predict and assess various engineering issues within the civil engineering domain [13-15]. In pavement engineering, ML's potential remains largely untapped due to the unpredictable nature of data. However, as ML advances, it could increasingly address challenges in this field, offering innovative solutions [16]. An insightful study by M. Asadi et al. showcased this potential, using ML to predict LAA values based on rock aggregate properties, such as uniaxial strength. The research revealed that specific ML methods outperformed traditional correlation approaches [17]. The research presented thus far has highlighted a notable gap in the existing literature regarding the influence of various factors on abrasion resistance, encompassing both the physical and mechanical properties of aggregates. The focus of much prior research was primarily on establishing linear correlations between the two factors and their mutual effects. However, the impact of multiple variables derived from available data on the abrasion resistance of various aggregates, as determined by MD abrasion tests, has not been comprehensively explored in previous studies. The MD test, conducted with water to mimic field conditions by exposing aggregates to weathering and mechanical stress, has been confirmed as reliable and repeatable by past studies and Department of Transportation (DOT) reports. Consequently, it is required for quality control of aggregates in projects, especially road construction.

In this paper, the impact of physical and mechanical properties of aggregates on their abrasion resistance based on the MD abrasion test was investigated using different ensemble ML methods. The durability and strength of aggregate were predicted by evaluating several physical and mechanical properties, including water absorption, specific gravity (measured in dry, saturated-surface-dry (SSD), and apparent states), magnesium sulfate soundness, alcohol-water freezing and thawing tests, and the LAA tests. A set of laboratory data was collected to construct ML algorithms. Python was employed to analyze different algorithms and predict the abrasion values based on various laboratory tests. Finally, to evaluate the accuracy and effectiveness of these machine learning models, four commonly employed performance metrics were used: Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the R-squared (R^2 score).

2. Materials and Methods

ML is widely applied in civil engineering for prediction, classification, and solving complex mathematical challenges. In the context of road safety, one critical attribute of asphalt pavement is its friction resistance. Insufficient friction between car tires and road surfaces is a significant contributor to traffic accidents [18]. Ensuring maximum skid resistance under diverse weather conditions are imperative. Primary factors influencing pavement friction includes aggregate type and pavement surface texture. Standard tests, including the vital MD test, evaluate the durability and abrasion resistance of aggregates. In order to achieve a good prediction of the abrasion resistance in aggregates by the MD test, consideration was given to various tests and aspects that might have an impact on aggregate abrasion, such as water absorption, specific gravity in dry, SSD, apparent conditions, and soundness tests of magnesium sulfate, along with alcohol-water freezing, thawing, and the LAA test. Experimental data from various studies was collected to formulate the model. Due to the use of labeled datasets, the model is defined by its ability to be trained by algorithms that accurately predict outcomes. The development of super learner models was undertaken by utilizing a varied ensemble of methods, including random forest (RF), gradient boosting machine (GBM), extreme gradient boosting (XGBoost), adaptive boosting (AdaBoost), categorical gradient boosting (CatBoost), and stacking by utilizing linear regression as the meta-learner to combine various algorithms for enhanced accuracy. The optimal conditions for each ensemble algorithm were obtained through a process of hyper-parameter optimization, utilizing a grid search method to achieve the best performance. For the assessment of the accuracy and effectiveness of these ML models, four widely used performance metrics, namely MSE, RMSE, MAE and the R^2 score, were utilized. These metrics provided valuable insights into the models' performance in accurately predicting the abrasion resistance of aggregates through the MD test and capturing the underlying patterns in the data. The MSE quantifies the average squared difference between the actual and predicted values, presenting an overall measure of prediction accuracy. Derived from MSE, the RMSE represents the square root of the average squared error, providing a measure of prediction deviation relative to the actual values. The MAE calculates the average absolute difference between the actual and predicted values, offering a straightforward measure of the model's predictive errors. Additionally, the R^2 score statistic serves as a crucial indicator of how well the influence of an independent variable explains the variance in a dependent variable. It aids in determining the extent to which the variability in the abrasion resistance of aggregates through the MD test can be attributed to the variations in the input features.

2.1. Ensemble Learning Methods

Ensemble learning methods can be categorized into three distinct groups: bagging, boosting, and stacking, as demonstrated in Fig. 1. In this study, all these methods were utilized. Firstly, the bagging method, specifically the RF algorithm, was employed. Secondly, the boosting method, encompassing GBM, XGBoost, AdaBoost and Catboost, was put to use. Additionally, a stacking ensemble model, capable of combining the strengths of both boosting and bagging methods was applied [19].

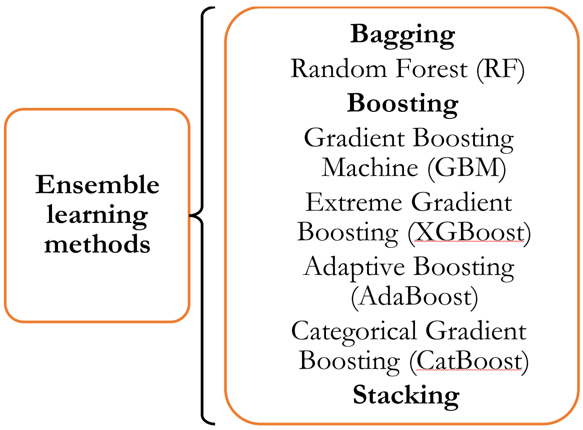

Fig. 1. Flowchart of ensemble methods.

2.1.1. Random FOREST (RF)

The RF is an ensemble learning method rooted in bagging, utilized for both regression and classification purposes. In the RF, trees operate independently in parallel, with no interaction among them during the construction phase. The introduction of random elements

DOI:10.4186/ej.2024.28.3.15

allows each tree to make independent predictions, even though they are developed using a deterministic algorithm and a fixed set of training data. [20].

2.1.2. Gradient boosting machine (GBM)

GBM is an ensemble machine learning technique that sequentially merges the forecasts of several basic models, often decision trees, to enhance overall predictive accuracy. By adjusting the model's weight based on errors, it systematically minimizes prediction errors, progressively enhancing model accuracy. GBM possesses the capability to uncover any nonlinear relationships between the target and features of the model, offering excellent usability that can handle missing values and outliers without requiring any special treatment. [21].

2.1.3. Extreme gradient boosting (XGBoost)

XGBoost is an optimized distributed Gradient Boosting Machine (GBM). It employs an ensemble learning approach that combines predictions from numerous weak models to generate a more robust prediction. An impressive aspect of XGBoost is its skillful management of missing values, which enhances its efficacy. This approach effectively addresses overfitting while also improving computational efficiency [22].

2.1.4. Adaptive boosting (AdaBoost)

The fundamental principle of AdaBoost involves iteratively adjusting parameters associated with a designated set of functions. AdaBoost's efficacy lies in its capability to transform a group of weak learners into a robust ensemble learner, often delivering remarkable predictive accuracy. Nonetheless, caution is advised against overfitting, particularly when the weak learners are overly complex [23].

2.1.5. Categorical gradient boosting (CatBoost)

CatBoost builds upon the principles of decision trees and gradient boosting. The main concept behind CatBoost involves sequentially integrating numerous weak models to create a robust predictive model through iterative improvement. This model utilizes the complete dataset during training and incorporates random permutations for each instance. It introduces a novel approach for computing leaf values during the selection of tree structure; through these advancements, CatBoost significantly improves model performance and the ability to generalize [24]. CatBoost diverges from typical gradient boosting models by employing oblivious trees. This approach dictates that trees are grown under the constraint that nodes at the same level must test the same predictor with the same condition.

2.1.6. Stacking

Stacking is widely recognized as one of the most utilized and effective ensemble techniques within the realm of ML as distinguished from bagging and boosting, stacking integrates numerous classifiers or regressors developed using various ML algorithms, functioning across distinct levels or layers [25]. Given the stacking ensemble model's potential to generate diverse permutations through various ML algorithms, this study focused on applying the super learner technique. Specifically, linear regression was employed as the meta-learner to combine different algorithms, with the objective of achieving improved accuracy. Figure 2 depicted the process of the stacking algorithm.

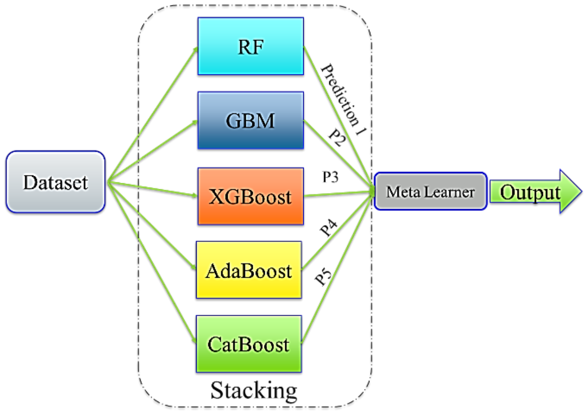

Fig. 2. Stacking algorithm process.

3. Data Collection and Processing

For MD prediction, data from various peer-reviewed publications, some of which are referenced in Table 1 were collected, comprising 300 entries with various physical and durability factors affecting aggregate abrasion. The dataset, derived from standard tests, included 7 input features: Los Angeles Abrasion (LAA), Magnesium Sulfate Soundness (MSS), Water–Alcohol Freezing and Thawing (AFT), Water Absorption (ABS), Bulk, Saturated Surface Dry (SSD), and apparent specific gravity (SGBULK, SGSSD, SGAPP). The output was MD abrasion. To ensure dataset quality, missing data and outliers were filtered, resulting in a final dataset of 282 records specifically focused on MD abrasion. The dataset's statistical characteristics are summarized in Table 2.

3.1. Micro-Deval Abrasion (MD)

The MD test represents one of the most important methods for quality control, evaluating the abrasion resistance and durability of mineral aggregates Fig. 3. This test was conducted according to ASTM D6928 and ASTM D7428. An aggregate sample is placed in a test container with abrasive balls and water. The container is then rotated at 100 rpm for a specific time. The quality of an aggregated report is based on how much mass loss occurred [2].

Table 1. Some sources of collected Data.

ID	Type	LAA	MSS	AFT	ABS	SG_{BULK}	SG_{SSD}	SG_{APP}	MD	Ref
1	Dolomite	31	12.2	7	1.6	2.58	2.63	2.7	17	[17]
2	Granite	29	23.9	5.75	2.2	2.51	2.57	2.66	22.2	[18]
3	Meramec	21	1.1	2.375	0.6	2.75	2.76	2.79	9.7	[26]
4	Limestone	26	3	2.8	0.7	2.67	2.69	2.72	11	[27]

Unit of LLA, MSS, AFT, ABS and MD (%)

Table. 2. Statistical properties of dataset.

	LAA	MSS	AFT	ABS	SG_{BULK}	SG_{SSD}	SG_{APP}	MD
count	282	282	282	282	282	282	282	282
mean	25.44	7.59	4.43	1.50	2.17	2.19	2.25	14.70
std	8.25	7.51	2.90	1.16	0.74	0.74	0.73	7.34
min	9.36	0.03	0.60	0.10	0.98	0.99	1.01	1.40
max	56.88	36.70	13.90	5.91	2.88	2.88	2.91	39.98

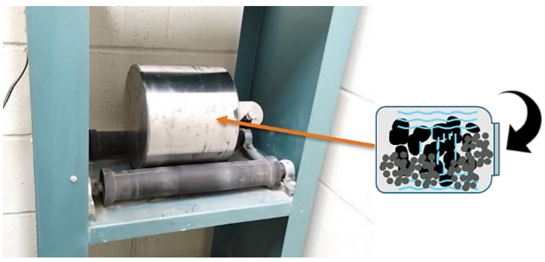

Fig. 3. MD apparatus.

Figure 4 displayed the Pearson correlation coefficient between various attributes. This coefficient revealed a weak and negative correlation among certain variables. While the Pearson correlation coefficient is commonly employed to assess the strength and direction of linear relationships between variables, it may not reliably identify nonlinear relationships.

Pair plots (Fig. 5) provided valuable insights into the data distribution, revealing patterns and potential anomalies. This data visualization tool facilitated a deeper understanding of the relationships between variables, guiding informed decisions during the analysis and modeling process. A data splitting strategy was employed to develop and assess ML models. The dataset was divided into two segments: a training sample, which comprised 80 percent of the data, and a test sample, which contained the remaining 20 percent. The training sample was utilized for the construction and training of the ML model using various super learner methods, aimed at uncovering the underlying patterns and relationships between the input features, such as LAA, MSS, AFT, ABS, SGBULK, SGSSD and SGAPP [18]. Once the model was trained, its performance was evaluated using the test sample, which had been withheld during the training phase. The ability to accurately predict the MD abrasion based on the provided input data was assessed by the trained model. Valuable insights into the model's generalization and its performance on previously unseen data were provided by this evaluation. Through this rigorous testing, it was ensured that new data could be effectively handled by our model, enabling accurate predictions in real-world scenarios.

3.2. Hyperparameter Tuning

Hyperparameter tuning constitutes a crucial step in the development of robust ML models. By tuning the ML model, overfitting can be mitigated, thus enhancing the model's adaptability to unseen data [25]. Optimal hyperparameter selection also plays a determinant role in augmenting model accuracy. Various approaches have been proposed to automate the selection of hyperparameters, such as grid search and random search hyperparameter optimization, aiming to avoid manual tuning. Grid search was utilized in this study, involving the construction of a model for each conceivable combination of the provided hyperparameter values. Each model was then evaluated, and the architecture yielding the most optimal results was selected.

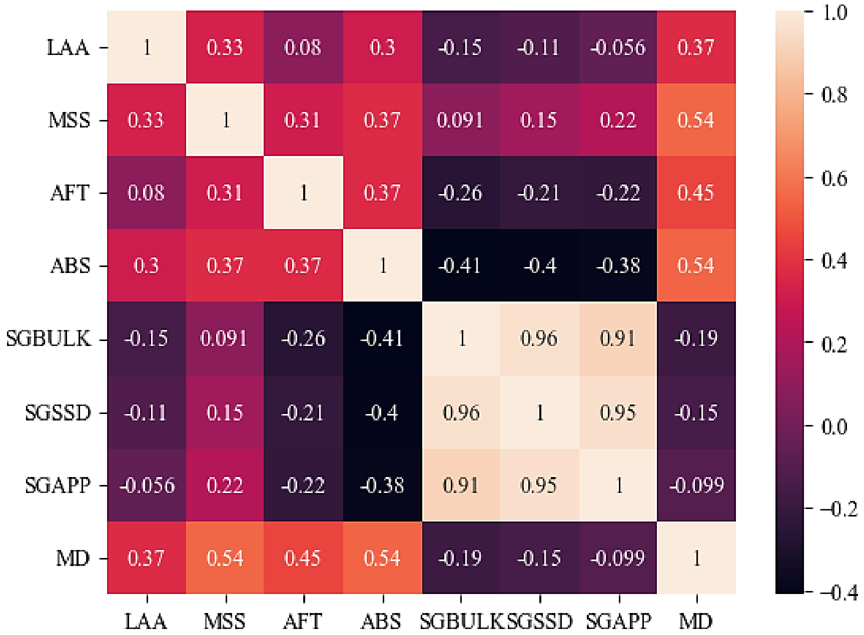

Fig. 4. Pearson correlation coefficient for the dataset.

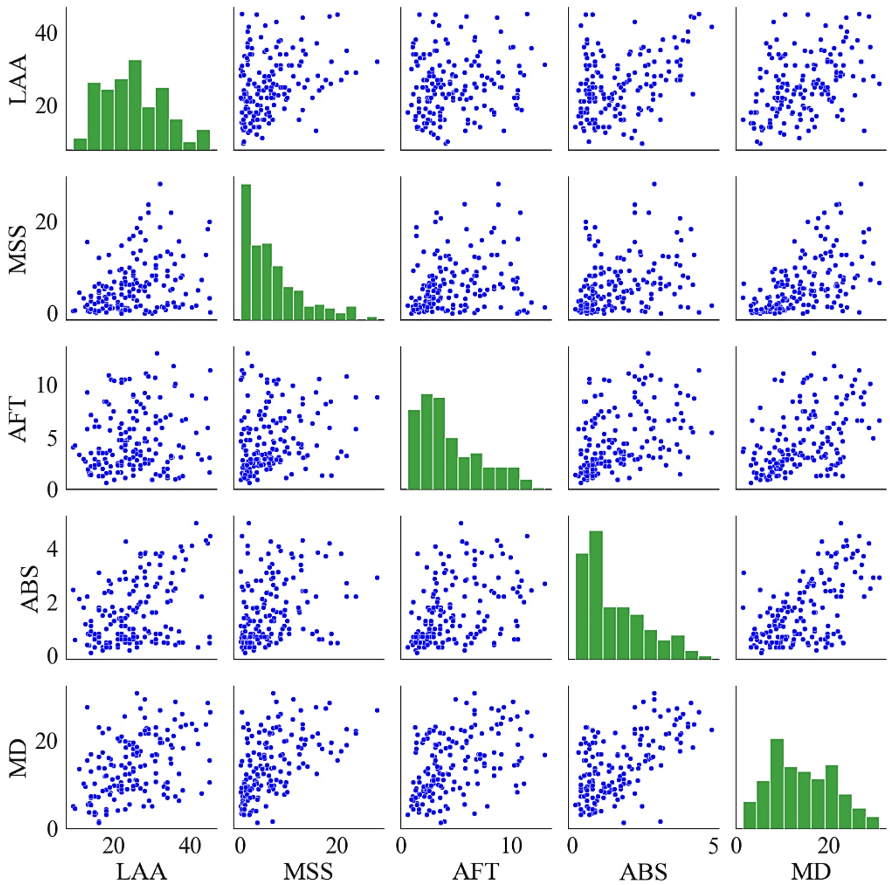

Fig. 5. Pair graph of dataset.

4. Results and Discussion

The development of accurate prediction models for estimating the abrasion resistance of aggregates through the MD test constituted the main objective of this study. In order to achieve this goal, supervised ML algorithms were employed, with a particular focus on various ensemble models. Six ensemble models, including RF, GBM, AdaBoost, XGBoost, CatBoost, and a stacking method combining elements of boosting and bagging, were implemented using the scikit-learn library in Python. The utilization of these super learner models aimed to predict the abrasion resistance of aggregates through the MD test, considering various physical and durability factors that could potentially influence aggregate abrasion. By training each model with the labeled dataset, intricate relationships between the input features and their corresponding MD abrasion resistance values were acquired.

4.1. Performance of ML Models

High accuracy in predicting the abrasion resistance of aggregates through the MD test was demonstrated by the Stacking, GBM, and CatBoost models, which achieved the highest prediction accuracy with R^2 score exceeding 0.94 for the stacking model and approximately 0.92 for both the GBM and CatBoost models, as depicted in Table 3 and Fig. 6. This outcome suggests that the relationship between the abrasion resistance of aggregates through the MD test and other variables is not primarily linear, owing to the intricacies of the dataset and the complex interplay of various factors. Notably, the dataset incorporates abrupt changes in specific values, which can adversely affect the accuracy of sensitive algorithms like Adaboost, resulting in a lower R^2 score and higher MSE, RMSE and MAE. Conversely, other algorithms displayed more accurate predictions on the test data, mainly due to their adeptness in capturing the nonlinear nature of the dataset.

The stacking model exhibited the best predictive performance, as indicated in Table 3, with the lowest MSE, RMSE, and MAE values for the test and train data, along with the highest coefficient of determination compared to that of the other models. The MAE and MSE and RMSE values were 0.99, 2.14 and 1.46 respectively, for the stacking ensemble model. It is evident that the model effectively captured the trend in the data, showcasing robust performance on both the train and test datasets.

Table 3. Super learner models metrics.

Model	MAE		MSE		RMSE		R^2 score	
	Test	Train	Test	Train	Test	Train	Test	Train
RF	1.74	1.71	5.85	5.44	2.41	2.33	0.85	0.84
GBM	1.22	1.27	3.22	3.23	1.79	1.8	0.92	0.92
AdaBoost	2.79	3.21	11.77	14.26	3.43	3.77	0.71	0.67
XGBoost	0.88	0.88	4.46	4.46	2.11	2.11	0.89	0.89
CatBoost	0.9	0.9	3.204	3.205	1.78	1.78	0.92	0.92
Stacking	0.99	0.99	2.14	2.14	1.46	1.46	0.948	0.947

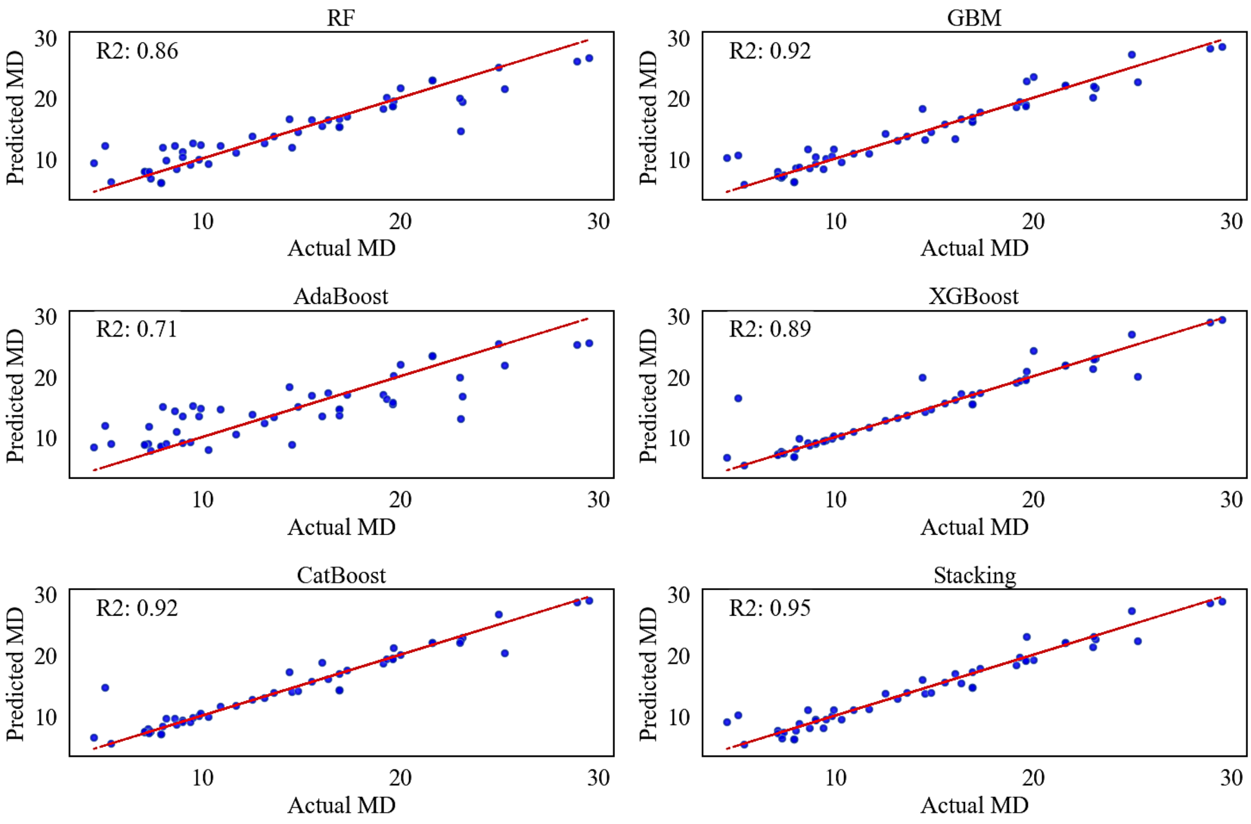

Fig. 6. Predicted vs Actual values of the MD for the different methods (test dataset).

DOI:10.4186/ej.2024.28.3.15

Both the CatBoost and GBM models demonstrated a high capacity to generalize learned patterns, as evidenced by the notably similar prediction errors observed between the training and testing sets, which supports the notion that transitioning from linear regression to ensemble models, particularly boosting methods, strengthens the capability to capture non-linear relationships inherent in the data. For the CatBoost model, the MAE, MSE, and RMSE values were 0.9, 3.2, and 1.78, respectively, and the GBM model produced MAE, MSE, and RMSE values of 1.22, 3.22, and 1.77, respectively. This enhancement consequently leads to a significant improvement in prediction accuracy for the specific task of estimating the abrasion resistance of aggregates through the MD test. Table 4 presented the predicted values of the MD by different methods utilized in this study, derived from randomly selected values from the test set and compared with their corresponding actual values. In most cases, the prediction results were considered acceptable.

The comparison between the actual and predicted values by the CatBoost model revealed a mean difference of 0.98, with a standard deviation of the difference equal to 1.59, showcasing the powerful performance of the model in predicting the MD. The MD abrasion values of aggregates are influenced by several factors, including water absorption, specific gravity, soundness tests, and LAA results. To assess the impact of each feature on the target, the feature importance for the CatBoost model, which exhibited the highest prediction accuracy and the lowest errors, was computed. The result of importance values for each feature on the target function were presented in Fig. 7 left) bar chart and right) radar plot, revealing that the MSS had the most significant effect on the predicted abrasion values by the MD. Subsequently, ABS and LAA also demonstrated notable importance, while other features exhibited relatively lower importance.

The feature importance analysis revealed that the abrasion values of aggregates are influenced by several factors, such as MSS, ABS, and LAA. As a result, it is recommended that these tests should be conducted on the aggregate to ensure the attainment of an accurate value of abrasion resistance. The influence of these factors on the abrasion values highlights the necessity of comprehensive testing to capture the intricate interplay of variables affecting the abrasion resistance of aggregates. Previous studies have indicated that the strength of rocks and aggregates is influenced by various geological features. Consequently, these research efforts have led to improve regression and analytical models where the assessment of polishing and abrasion properties is based on specific aggregate types or groups that share similar geological characteristics [28-29]. However, the feature importance analysis in this study revealed that the impact of multiple variables derived from available data on the abrasion resistance of various aggregates could be effectively assessed using different ML methods.

Table 4. Random selected prediction values.

Actual Value	Predicted Value				
MD	RF	GBM	AdaBoost	XGBoost	CatBoost
12.5	12.96	14.13	14	12.68	12.59
25.25	21.99	22.58	21.5	19.92	20.32
19.1	18.51	18.43	17.27	19.02	18.59
20	21.93	23.4	21.42	24.19	19.93
19.3	20.04	19.35	16.74	19.26	19.24
16.32	16.74	16.52	16.65	17.13	16.04

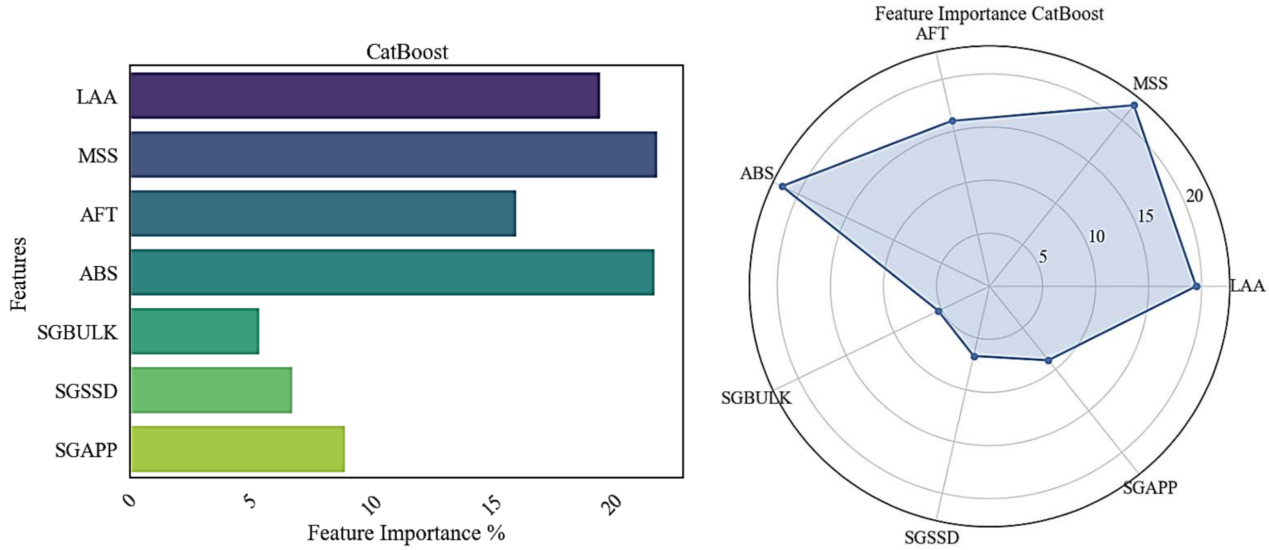

Fig. 7. Feature importance of CatBoost model Left) bar chart, Right) radar plot.

5. Conclusions

The primary objective of this research was to use real data acquired from various standard tests on diverse aggregates to predict the degree of abrasion of the aggregate obtained from the MD test. Certain durability and physical tests, such as the LAA test and soundness tests, have been known to be time-consuming, destructive, and less reliable. Consequently, the growing use of the MD is attributed to its capability to simulate field conditions and yield valid results. Additionally, the escalating extraction of mineral aggregates for construction projects has led to an increasing trend of utilizing recycled materials. In this context, the development of ML methods has been proven to be advantageous in predicting required values obtained through destructive tests, thereby contributing to the preservation of aggregate materials.

This study has demonstrated how ML, particularly ensemble methods, can be effectively employed in the field of pavement engineering, facilitating precise predictions. Previously, a significant gap existed between pavement engineering and ML due to insufficient laboratory data and diverse features, leading most researchers to rely on simple correlations between data. However, with the application of ML, it has now become feasible to assess models with a broader array of features, utilizing advanced methods and algorithms.

The exceptional precision demonstrated in predicting the abrasion resistance of aggregates through the MD test was underscored by the provided R^2 score values, with the stacking, GBM, and CatBoost models exhibiting an impressive R^2 score of 0.95 for the stacking model and approximately 0.92 for both the GBM and CatBoost models.

The capacity of the CatBoost and GBM models to generalize learned patterns was demonstrated, as confirmed by the remarkably similar prediction errors observed between the training and testing sets, reinforcing the concept that transitioning from linear regression to ensemble models, especially boosting methods, enhances the ability to capture non-linear relationships inherent in the data. The analysis of the comparison between the actual and predicted values by the CatBoost model uncovered a mean difference of 0.98, accompanied by a standard deviation of the difference equaling 1.59. These results stand as a testament to the formidable performance of the model in accurately predicting the MD.

The feature importance analysis uncovered that the abrasion values of aggregates are influenced by several factors, such as MSS, ABS, and LAA. Therefore, the utilization of multiple test methods could potentially yield a more reliable assessment of aggregate durability and abrasion.

Acknowledgement

The authors would like to express their heartfelt gratitude to Missouri University of Science and Technology for their invaluable support and assistance throughout this research.
The dataset used in this study was collected from various peer-reviewed publications and reports, as referenced.

References

[1] I. Ugur, S. Demirdag, H. Yavuz, "Effect of rock properties on the Los Angeles abrasion and impact test characteristics of the aggregates," *Materials Characterization*, vol. 61, no. 1, pp. 90–96, 2010.

[2] *ASTM Standard Test Method for Resistance of Coarse Aggregate to Degradation by Abrasion in the Micro-Deval Apparatus*, ASTM D6928-17.

[3] *ASTM Standard Test Method for Resistance to Degradation of Small-Size Coarse Aggregate by Abrasion and Impact in the Los Angeles Machine*, ASTM C131/C131M-20.

[4] J. L. Zhang and M. C. He, "Discussion on the causes of asphalt concrete pavement loose damage and preventive measures," *Highway*, vol. 10, pp. 225–227, 2009.

[5] A. Roshan and A. Ghasemi, "Laboratory investigation of micronized lomashell powder effects on asphalt binder and mix performance," *Journal of Engineering & Technological Sciences*, vol. 55, no. 3, pp. 324-339, 2023.

[6] S. A. Senior and C. A. Rogers, "Laboratory tests for predicting coarse aggregate performance in Ontario," in *Transportation Research Record 1301*. Transportation Research Board, National Academy of Sciences, pp. 97-106.

[7] L. A. Cooley Jr. and R.S. James, "Micro-deval testing of aggregates in the southeast," *Transportation Research Record*, vol. 1837, no. 1, pp. 73-79, 2003.

[8] M. Sansalone, J. M. Lin, and W. B. Streett, "Determining the depths of surface-opening cracks using impact-generated stress waves and time-of-flight techniques," *ACI Materials Journal*, vol. 95, pp. 168–177, 2018.

[9] M. Takarli, A. Mardele, and W. Prince, "Adaptation of the micro-deval test for assessing the surface weathering degree of granite," *Journal of ASTM International*, vol. 6, no. 9, pp. 1-1, 2009.

[10] S. Adomako, C. J. Engelsen, R. T. Thorstensen, and D. M. Barbieri, "Review of the relationship between aggregates geology and Los Angeles and micro-Deval tests," *Bulletin of Engineering Geology and the Environment*, vol. 80, no. 3, pp. 1963-1980, March 2021.

[11] J. M. Palomino Ojeda, S. Rosario Bocanegra, and L. Quiñones Huatangari, "Determination of the compressive strength of concrete using artificial neural network," *Int. J. Eng. Technol. Innov.*, vol. 11, no. 3, pp. 204–215, 2021.

[12] I. Deiros, C. Voivret, G. Combe, and F. Emeriault, "Quantifying degradation of railway ballast using numerical simulations of micro-Deval test and in-situ conditions," *Procedia Engineering*, vol. 143, pp. 1016-1023, 2016.

[13] R. Biswas, B. Rai, P. Samui, and S. S. Roy, "Estimating concrete compressive strength using MARS, LSSVM and GP," *Engineering Journal*, vol. 24, no. 2, pp. 41-52, Mar. 2020.

[14] M. Adamu, P. Trabanpruek, V. Limwibul, P. Jongvivatsakul, M. Iwanami, and S. Likitlersuang, "Compressive behavior and durability performance of high-volume fly-ash concrete with plastic waste and graphene nanoplatelets by using response-surface methodology," *Journal of Materials in Civil Engineering*, vol. 34, no. 9, p. 04022222, Sep. 2022.

[15] H. W. Iqbal, K. Hamcumpai, P. Nuaklong, P. Jongvivatsakul, S. Likitlersuang, C. Chintanapakdee, and A. C. Wijeyewickrema, "Effect of graphene nanoplatelets on engineering properties of fly ash-based geopolymer concrete containing crumb rubber and its optimization using response surface methodology," *Journal of Building Engineering*, vol. 75, p. 107024, Sep. 2023.

[16] S. O. Aregbesola and Y. H. Byun, "Classification of geogrid reinforcement in aggregate using machine learning techniques," *International Journal of Geo-Engineering*, vol. 15, no. 1, p. 4, Feb. 2024.

[17] M. Asadi, A. TaghaviGhalesari, and S. Kumar, "Machine learning techniques for estimation of Los Angeles abrasion value of rock aggregates," *European Journal of Environmental and Civil Engineering*, vol. 26, no. 3, pp. 964-977, 2022.

[18] R. E. Weyers, "Testing methods to determine long-term durability of Wisconsin aggregate resources," Wisconsin Highway Research Program, 2005.

[19] A. Roshan and M. Abdelrahman, "Predicting flexural-creep stiffness in bending beam rheometer (BBR) experiments using advanced super learner machine learning techniques," *Res. Eng. Struct. Mater.*, to be published.

[20] A. Upadhya, M. S. Thakur, N. Sharma, and P. Sihag, "Assessment of soft computing-based techniques for the prediction of Marshall stability of asphalt concrete reinforced with glass fiber," *International Journal of Pavement Research and Technology*, vol. 15, no. 6, pp. 1366-1385, 2022.

[21] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, vol. 1189, no. 232, Oct. 2001.

[22] J. Xie, J. Huang, and D. V. Griffiths, "Learning from prior geological information for geotechnical soil stratification with tree-based methods," *Engineering Geology*, vol. 327, p. 107366, Dec. 2023.

[23] S. Demir and E. K. Sahin, "An investigation of feature selection methods for soil liquefaction prediction based on tree-based ensemble algorithms using AdaBoost, gradient boosting, and XGBoost," *Neural Computing and Applications*, vol. 35, no. 4, pp. 3173-3190, Feb. 2023.

[24] F. Zhang and H. Fleyeh, "Short-term electricity spot price forecasting using CatBoost and bidirectional long short-term memory neural network," in *2019 16th International Conference on the European Energy Market (EEM)*, Sep. 2019, pp. 1-6.

[25] D. H. Wolpert, "Stacked generalization," *Neural Networks*, vol. 5, no. 2, pp. 241-259, Jan. 1992.

[26] E. Deef-Allah, K. Broaddus, and M. Abdelrahman, "Evaluation of alternatives to calcined bauxite for use in high friction surface treatment (HFST) in Missouri," Construction and Materials Division, Missouri Department of Transportation, Sep. 2021.

[27] P. W. Jayawickrama, S. Hossain, and A. R. Hoare, "Long-term research on bituminous coarse aggregate: Use of micro-Deval test for project level aggregate quality control," Texas Transportation Institute, The Texas A&M University System College Station, Rep. no. FHWA/TX-06/0-1707-9, Mar. 2007.

[28] M. Kamani and R. Ajalloeian, "Evaluation of the mechanical degradation of carbonate aggregate by rock strength tests," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 11, no. 1, pp. 121-134, Feb. 2019.

[29] L. O. Afolagboye, A. O. Talabi, and O. O. Akinola, "Evaluation of selected basement complex rocks from Ado-Ekiti, SW Nigeria, as source of road construction aggregates," *Bulletin of Engineering Geology and the Environment*, vol. 75, pp. 853-865, May 2016.

Alireza Roshan, photograph and biography not available at the time of publication.

Magdy Abdelrahman, photograph and biography not available at the time of publication.