Performance Evaluation of ANN and Ensemble Learning Methods in Predicting Wear Properties of Porcelain Ceramic Composites

Ahmet Gürkan YÜKSEK[a, *]

[a] SCU Artificial Intelligence and Data Science Application and Research Center, University of Sivas Cumhuriyet, Sivas, 58140, Türkiye; SCU Nanophotonic Application Research Center, Sivas, Türkiye

Corresponding Author: Tel:+90-(505)-781-18-01 E-mail address:agyuksek@cumhuriyet.edu.tr

Abstract. In this study, the fabrication and wear properties of aluminum titanate (Al_2TiO_5) and mullite ($3Al_2O_3$-$2SiO_2$) doped porcelain ceramic-composites produced by powder metallurgy method were extensively investigated. Porcelain ceramics were prepared by powder metallurgy and wear resistance and other mechanical properties were evaluated based on the data obtained. The wear results were modeled and analyzed using ensemble learning (EL) methods and Artificial Neural Networks (ANN). Among the ensemble methods, Boosting and random forest (RF) algorithms were used. Basic statistical measures such as R^2, RMSE, MAE and MAPE were used to evaluate model performance. Boosting and RF methods also gave good results, but ANN was found to be best in terms of accuracy and overall performance. In the study, pure porcelain (P), mullite doped porcelain (PM), aluminum titanate doped porcelain (PAT) and aluminum titanate-mullite doped porcelain (PMAT) porcelain models were investigated and compared separately. The results provide valuable contributions to the development of high-performance ceramic-composites in materials engineering and optimization of the wear behavior of these materials. This paper examines in detail the applicability and benefits of advanced machine learning methods in materials science.

Keywords. Porcelain, Wear Resistance, Artificial Neural Networks, Ensemble Learning.

1. Introduction

Porcelain is a ceramic material whose microstructure is fine-grained and non-porous, with high strength and generally translucent properties. It is highly vitrified and characteristically white in color, offering superior mechanical and aesthetic properties. Porcelain consists mainly of kaolin, quartz and feldspar, fired at high temperatures to give the ceramic its characteristic properties. The triple composition ratios are typically 50% clay, 25% flux and 25% filler, which optimizes the mechanical and aesthetic properties of porcelain. Sintered between 1200°C and 1400°C to form a glass-ceramic composite, this ceramic material is well known for its semi-transparent properties. The sintering process, which is carried out at these high temperatures, creates the complex structure of porcelain and gives it superior mechanical strength and chemical resistance[1-5].

Aluminum Titanate (Al_2TiO_5) is known for its exceptional thermal shock resistance, which makes it resistant to sudden temperature changes. It also has low thermal conductivity properties, thus efficiently insulating heat in high temperature applications, and exhibits chemical resistance in molten metals, making it ideal for corrosive environments. Al_2TiO_5 maintains dimensional stability under high temperature changes due to its low coefficient of thermal expansion. These properties make Al_2TiO_5 an excellent material for industrial applications requiring high performance such as glass manufacturing, automotive parts and thermal processing equipment[2,6,7]. Mullite ($3Al_2O_3.2SiO_2$) exists as the only stable interphase in the Al_2O_3 - SiO_2 binary system, which makes it unique. Mullite also exhibits excellent thermal shock resistance due to its thermal and chemical stability and low coefficient of thermal expansion [8-11]. Due to their high hardness and low coefficient of friction, ceramic materials can markedly enhance the performance of components, particularly in contact-load applications. The wear behavior of ceramic materials is closely related to their microstructural properties and crystal structure. The fact that ceramics exhibit anisotropic properties directly affects the wear resistance of these materials. In addition, the high temperature resistance of ceramics is also related to the material's thermal expansion coefficient and thermal shock resistance. The anisotropic nature of their crystal structure and microstructural features determines the wear and thermal resistance performances of these materials, which makes them reliable even under extreme conditions[12-18].

ML extracts information from inputs through iterative learning processes and aims to improve the accuracy of models [19,20]. Algorithms that learn from data automate various tasks by identifying and analyzing patterns in the inputs. It has been successfully used in predictive analyses in engineering and other fields[21]. It allows to understand complex structures in data and to perform operations on these structures. Pattern recognition is one of the important applications of ML and aims to automate processes [22-24]. Feature selection and algorithm selection are important aspects of predictive analysis, and the use of ML has been instrumental in advancing these areas [25]. A comprehensive review of ML algorithms has highlighted their ability to train with data and make predictions without explicit human guidance [26]. EL is a technique where multiple models or algorithms can be combined for a better and more accurate prediction model[27, 28]. The basic principle behind the method is that by combining predictions from multiple models, the overall prediction is more reliable and robust than a single model[29]. EL can improve the performance of ML algorithms by reducing bias, increasing stability and improving generalization[30]. In general, EL is a powerful tool in the field of ML that leverages the strengths of multiple models to improve prediction accuracy and robustness[31]. EL offers various techniques in prediction methods. Bagging provides a way to combine predictions from models developed on different subsets[32], while Boosting focuses on incorrect predictions to turn weak learners into strong predictors[33]. Random Forest (RF) improves accuracy by combining the predictions of many decision trees[34]. These methods represent various aspects of EL and can be used in different scenarios to improve prediction performance.

In engineering applications, ensemble methods are employed to enhance the accuracy of predictions, augment the capacity for generalization, and circumvent the issue of overfitting[35]. Combining different models enhances generalization, prevents overfitting, and provides more reliable forecasts. Ensemble methods offer a comprehensive prediction by integrating diverse models, aiding in more robust decision-making and error compensation[36]. To illustrate, in the case of Bagging, models trained with random subsets of the data provide more stable results by reducing variance[37]. In contrast, in Boosting, each model gradually corrects errors by focusing on areas where the previous model is weak[38]. Ensemble methods provide more reliable and stable forecasts by increasing generalizability while reducing bias and variance in forecasts[39]. The stability achieved by such methods is dependent on the strategy employed, which may include majority voting, averaging, or weighted aggregation[40]. The objective is to balance the biases and uncertainties of individual models, thereby producing a more accurate and reliable result[41]. In the field of materials science and engineering, ensemble methods play a crucial role in predicting and analyzing material properties. Combining multiple machine learning models to simulate complex behaviors linked to microstructure reduces errors and enhances generalization capability. Methods such as Boosting and RF improve experimental accuracy while preventing overfitting, making predictions more reliable under real-world conditions. These approaches are critical tools for optimizing material performance and developing new composite materials[42].

In this study, by combining materials engineering and artificial intelligence techniques, the fabrication and wear properties of aluminium titanate (Al_2TiO_5) and mullite (M, $3Al_2O_3$-$2SiO_2$) doped porcelain ceramic-composites were investigated in depth. These ceramic-composites, produced by powder metallurgy, are characterised by their superior mechanical and thermal properties. The experimental wear results were modelled and analysed using advanced machine learning methods such as EL and ANN. The preparation of porcelain ceramics was carried out by powder metallurgy and based on the data obtained, wear resistance and other mechanical properties were evaluated. This study provides important insights into how the results obtained by the integration of innovative manufacturing techniques and artificial intelligence-based modelling methods in materials engineering can be used in engineering applications[43-46].

2. **Methods and Theories**

This study outlines a methodology for modeling regression problems using ensemble learning, consisting of three phases: data collection and preparation, model selection and training, and model evaluation. Initially, a suitable dataset is collected and cleaned. Various base models are selected and trained on this dataset. Ensemble methods such as bagging, boosting, or random forest are then used to combine predictions from these base models. The ensemble model is

subsequently evaluated on a test set, with performance metrics measured and adjustments made if necessary. Finally, the results are reported and interpreted, systematically guiding the application of ensemble learning to regression problems, as illustrated in Figure-1. Finally, cross-validation is used to increase the generalizability of the model and reduce the risk of overfitting[47]. These steps leverage EL's strengths in accuracy, generalization, and robustness to solve regression problems.

Figure 1: General Process Design of the Model

Experimental data were collected to investigate the wear behavior of porcelain ceramics. Firstly, porcelain samples were prepared from basic components such as kaolin, quartz, and feldspar. Al_2TiO_5 and mullite ($3Al_2O_3$-$2SiO_2$) additives were added to the porcelain mixture. The prepared mixtures were sintered at temperatures between 1100°C and 1200°C using the conventional powder metallurgy method. In this study, the production and wear properties of porcelain-ceramic composites with aluminum titanate and mullite additives produced by powder metallurgy method were investigated and then the experimental wear data obtained were analyzed. The data obtained were modeled using EL and ANN. Powder metallurgy method was used for the preparation of porcelain ceramics. AT, Al_2TiO_5 and mullite (M, $3Al_2O_3.2SiO_2$) ceramic powders were achieved through the reaction sintering of Al_2O_3, SiO_2 and TiO_2 powders at 1550°C and 1400°C for 2 hours, respectively. After crushing, grinding, and sieving, aluminum titanate and mullite powders were made ready for use. These powders were blended with porcelain (P) in weight percentages of 0% and 20%. Porcelain ceramics doped with AT and M were fabricated using the powder metallurgy method. The blends were prepared by mechanical alloying in alumina ball mills, shaped in a dry press and then sintered under normal atmospheric conditions. The mixture powders were formed into 56x12x10 mm preforms by uniaxial pressing at 200 MPa. The green compacts were air sintered in a high temperature furnace (Protherm™ Furnace) at 1100-1200°C at a heating rate of 5°C/min for 1-5 hours.

Ceramic wear tests were carried out using a Plint brand abrasion tester. A steel disk was preferred as the wear disk in the tests. Abrasion tests were performed on each specimen for 5, 10, 15 and 20 minutes at 70, 90 and 120 N forces. After the initial weighing of the samples with a 0.0001 g precision balance, the wear amounts were calculated by weighing them again after the specified wear times [1,2]. Subsequent characterization studies of the sintered samples were performed, and the wear test results were transformed into data suitable for modeling using EL and ANN.

Table 1: Data Set Parametric Structure

The experimental data and input parameters summarized in Table-1 were used to develop regression models for the calculation of wear volume values. Figure-2 shows that different wear volume results were obtained for each measurement type (P,PM,PAT and PMAT). The P measurement results represent the data obtained from the basic porcelain composition, while the PM measurement results show the mullite-doped variant of porcelain. PAT measurement results reflect the wear performance of aluminum titanate-doped porcelain samples, while PMAT measurement results include data from both mullite and aluminum titanate-doped porcelain. The artificial intelligence models to be developed will focus on predicting the wear volume (mm^3) value using the input parameters wear force (N), wear time (min), fired temperature (°C) and fired time (hour). The test conditions and specimen compositions employed in experimental measurements exert a significant influence on the wear volume. Consequently, it is imperative that these input parameters are meticulously selected and processed for machine learning models to make accurate and reliable predictions. The models will be designed to predict the wear behavior of porcelain ceramics under different conditions by learning the effects of input parameters on the wear volume.

Figure 2: Wear Behavior of Porcelain-Ceramic Specimens

In this study, experimental data on the wear behavior of porcelain ceramics are trained using ensemble methods and analyzed in comparison with ANN approach. The aim is to train models, compare their performance, and identify the best approach to predict wear volume, providing a basis for future predictions. Additionally, the study compares the performance of ensemble methods with ANN using different training and learning processes. Boosting is a technique designed to enhance model accuracy

by iteratively focusing on correcting errors made by previous models in a sequence. Each new model in the boosting process is trained to address the weaknesses of its predecessors, which enables it to capture general patterns effectively. However, boosting can be time-intensive due to its sequential nature, requires careful hyperparameter tuning, and is sensitive to noisy data, potentially leading to overfitting if not managed properly. Bagging (Bootstrap Aggregating), on the other hand, builds multiple models independently by training each on a random subset of the data (with replacement). The predictions from these models are then aggregated, usually by averaging (for regression) or voting (for classification), to improve generalization and reduce overfitting. Bagging leverages diversity among models to stabilize predictions, but it demands significant computational resources, and its performance heavily relies on the diversity and quality of the data subsets used. The key difference between boosting and bagging lies in how the models are trained and aggregated. Boosting emphasizes learning from mistakes in a sequential and dependent manner, making it sensitive but highly effective for capturing patterns. Bagging, in contrast, trains models independently in parallel, focusing on reducing variance by averaging predictions from diverse models. RF, a bagging-based method, builds multiple decision trees on different data subsets and combines their outputs to achieve high accuracy while mitigating overfitting. Additionally, it enhances model interpretability by providing insights into feature importance. However, like bagging, it requires substantial computational resources and can be slow in model estimation. ANNs offer a flexible and robust framework for learning from large and complex datasets. They excel in discovering intricate patterns but require significant computational power, careful hyperparameter tuning, and sufficient data to prevent overfitting. Given these characteristics, the choice between boosting, bagging, and other methods like RF or ANN should align with the application's goals, dataset size, and computational constraints[20,27,33,34,47].

Ensemble boosting is a powerful learning method, at least for the accuracy of the predictive model. Boosting combines weak learners sequentially to build a strong learner that performs better than each individual model. Boosting works by iteratively adjusting the weights of training data points, focusing on those that were previously mispredicted. The key steps in boosting include, initialize weights, train weak learner, compute error, update weights and combine weak learners. Mathematically, the boosting algorithm can be described as Table-2.

Table 2: The Detail of Boosting

Where N is the number of training examples, I is the indicator function. Boosting enhances predictive accuracy by sequentially focusing on misclassified instances to reduce bias, improving model performance by combining multiple weak learners into a strong model, and being highly adaptable to various types of predictive tasks[48-51].

Ensemble Bagging, an acronym for Bootstrap Aggregating, is a machine learning method that aims to improve the accuracy and stability of models. Multiple models are trained on random subsets of the training data to prevent overfitting and reduce variance. The predictions of these models are usually averaged and combined to obtain more stable results. Bagging is particularly effective in regression tasks. Its basic steps are to create bootstrap samples, train independent models on these samples and combine the predictions to obtain the final result. This method improves generalization performance and minimizes errors. To illustrate the mathematical details further, let’s define each step more rigorously in Table3:

Table 3: The Detail of Bagging

Bagging is a robust technique for improving regression models in material engineering. By combining multiple models trained on different subsets of the data, bagging algorithms can effectively capture the complex patterns in data. Bagging enhances predictive accuracy by reducing variance through averaging multiple models, improving stability by training on different data samples to avoid overfitting, and allowing parallel training of models for computational efficiency[32,33].

Bagging and boosting are ensemble methods that aim to improve model performance but take different approaches. Bagging allows models to be trained on random subsets and the final prediction is obtained by the average or majority vote of these models. This method avoids overlearning by reducing variance but does not correct individual errors. Boosting trains models sequentially, allowing each model to learn

from the errors of the previous one. It reduces both bias and variance by giving more weight to misclassified examples. Although boosting produces robust results, it can be prone to overlearning in noisy data. In summary, while bagging reduces variance, boosting offers a more comprehensive improvement by reducing both variance and bias.

Random Forest (RF) is an effective ensemble method for classification and regression problems. It generates a large number of decision trees during training and combines the predictions of these trees to increase accuracy and reduce overfitting. RF is especially successful with high dimensional data and complex interactions between features. Each tree is generated using a random subset of the training data and selected features. The final prediction is made by majority voting in classification and by averaging the predictions in regression. RF reduces overall errors by bootstrap sampling and random feature selection. It also evaluates feature importance and identifies the variables that are most effective in predicting the target variable. The Random Forest (RF) algorithm is a machine learning algorithm that utilizes the parallel processing capability to efficiently process large datasets and generate reliable predictions. The algorithm's design incorporates features that enhance the accuracy and generalisability of models in high-dimensional datasets. Additionally, the diversity present in the decision trees mitigates the risk of overfitting, thereby improving the overall performance of the model. The basics of the RF algorithm can be described as Table 4:

Table 4: The Detail of RF

RF is a robust and effective EL method for regression and classification tasks. By combining multiple decision trees, it captures complex patterns in data, improves prediction accuracy, and provides insights into feature importance [34,51-55].

ANN model is an approach that aims to imitate human intelligence with computer systems. Based on this approach, it uses advanced algorithms and models to recognize patterns in data while performing problem solving tasks and aims to solve real world problems[56]. ANNs model the working principle of the human brain and process information by modelling the relationship of neurons in the human brain through mathematical methods available in its infrastructure[57]. In general, artificial neural network components: Input layer is the layer where data from the outside world enters the system. Hidden layers are the layers between the input layer and the output layer where information processing takes place. There can be one or more hidden layers and each layer can contain more than one artificial neuron. Output layer is the last layer of the artificial neural network and is usually the layer where results are produced to solve a specific problem[58,59]. The learning process of the ANN is achieved through the adjustment of weights and biases, with the objective of minimizing the discrepancy between the input data and the actual output. This process is executed through the utilization of the backpropagation algorithm, which updates the weights in each layer through error backpropagation[60]. The forward propagation and back propagation processes of ANN can be expressed in Table-5:

Table 5: The Detail of ANN

Here $z^{(l)}$, $a^{(l)}$, $w^{(l)}$, and $b^{(l)}$ represent the sum of weights, activations, weights, and biases in the first layer, respectively. $f(\cdot)$ is the activation function, $f'(\cdot)$ is the derivative of the activation function, $\nabla_a C$ is the derivative of the cost function with respect to the activations, and $\odot$ is the element-wise multiplication. The back-propagation algorithm is realized by calculating the error in the output of the network and the gradient of the weights and biases.

2.1. Process Algorithms of AI Models

Each of the three distinct EL and ANN methodologies entails a unique set of computational and information processing operations. These steps can be summarized as follows: data is obtained and processed, parametric operations are performed, results are generated, evaluated, and tested, and finally, the final model is built. All the processes are performed by means of computer software algorithms. The following section presents the pseudo-codes of the methods employed in the study, with the aim of illustrating the procedural or algorithmic differences inherent to the processes. The Boosting method offers an approach that trains weak learners sequentially, correcting the errors of the model at each step. Misclassified examples are given greater weight, and the new model attempts to rectify these errors (Table 6.a). The bagging method enhances the generalization of the model and circumvents overlearning

by utilizing multiple models. Each model is trained on a distinct subset derived through random resampling from the original dataset (Table 6.b). In the Random Forest (RF) method, multiple decision trees are integrated. Each tree is trained on a random subset of the data and a random subset of features at each discrimination point. This approach enhances performance and mitigates overlearning (Table 6.c). ANN models offer flexibility and powerful learning capabilities for large and complex datasets, they learn non-linear patterns in data thanks to their multi-layered structure (Table 6.d) [61].

Table 6: Pseudo Code for Artificial Intelligence Models [62].

2.2. Statistical Metrics for Comparison of AI Models

To evaluate the performance of artificial intelligence models, statistical measures such as R^2, Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are often employed. These measures provide disparate information regarding the predictive capacity and limitations of the model. It is crucial to comprehend the mathematical formulae, definitions, and advantages of these metrics to select the optimal evaluation method for the characteristics of the data and the objectives of the analysis. The combination of these metrics offers a more comprehensive view of model performance, facilitating a more robust evaluation[63-66].

R^2 is more commonly known as the coefficient of determination, and it represents how much of the variation in the dependent variable can be predicted by the independent variables. It has a value between 0 and 1, where 1 is perfect prediction **(Equ.1).**

$$1-\frac{\sum(y_i-y_i^{'})^2}{\sum(y_i-\hat{y}_i)^2} \qquad \text{Equation 1}$$

RMSE is a statistical measure that quantifies the average magnitude of the errors in a set of predictions, without considering their direction. It is calculated as the average of the absolute differences between the predictions and the actual observations over the test sample**(Equ.2).**

$$\sqrt{\frac{\sum(y_i-y_i^{'})^2}{n}} \qquad \text{Equation 2}$$

MAE is a statistical measure that quantifies the average magnitude of errors in a set of predictions. It does not consider the direction of the errors. The MAE is calculated as the average of the absolute differences between the predictions and the actual observations over the test sample**(Equ.3).**

$$\frac{\sum|y_i-y_i^{'}|}{n} \qquad \text{Equation 3}$$

MAPE is a percentage metric that expresses accuracy. It measures the average magnitude of the errors in a set of predictions, normalized by the actual values**(Equ.4).**

$$\frac{100\%}{n}\sum\left|\frac{y_i-y_i^{'}}{y_i}\right| \qquad \text{Equation 4}$$

3. Developed Method and Application

Experimental data were used for training the models. However, to prevent the models from being adversely affected by undesirable issues such as overfitting and memorization stemming from interactions with the training set, the dataset was initially divided into two parts. 75% of the dataset was randomly assigned as the model training set, while the remaining 25% was used for testing the models. The training set was utilized for training various artificial intelligence models, determining parameters, and constructing the final models. An important step in such analyses is testing and validating the performance of the models. Real-world data is often used in the testing and validation processes of the proposed systems. The key approach here is to ensure that the test data is not exposed to the models during training, thus preventing the models from developing any sensitivity to this data. The 25% of the data designated as the Model Test Set was prepared for this purpose and excluded from the training processes. After the models were trained with the "Model Training Set," they were tested using the "Model Test Set." The performance of the models was evaluated using fundamental statistical

comparison metrics commonly applied in regression analysis. Based on the results obtained from the test sets, the success of the developed models across various metrics was determined, thereby revealing their performance. Implementing this approach provides an objective evaluation of the models' performance on previously unseen data, which is crucial for validating their generalizability and robustness.

A series of machine learning models were trained using Boosting, Bagging, RF and Artificial Neural Network (ANN) models to predict the wear behavior of porcelain ceramics. During this training process, we took great care to determine the parameters, architectures and specific structures of each model. To optimize the performance of the models, we made various hyperparameter adjustments and selected the model that produced the best results by considering the statistical parameters. To assess the performance of the models, statistical metrics were used to evaluate the accuracy and reliability of the predictions. Once the training process was complete, the initial 25% test set was presented to each model. The results produced by the models were recorded separately and compared with the actual values. Figure-3 visually compares the prediction results of each model with the actual results, highlighting the outputs produced by the trained models when test data is presented alongside the actual outputs. These visualizations play an important role in comparing the success of the models.

Figure 3-Comparison of Model Predictions and Actual Outputs on Test Data

The results of each model are placed on the same graph and the behaviors of the models are presented in a visual comparison. To evaluate the effectiveness of artificial intelligence methods, various test metrics were used to evaluate the degree of agreement between the trained outputs of each method and the actual results. In this study, training was performed using Boosting, Bagging, RF and ANN models. At the end of the training process, the test set was subjected to these four models and the performance of each model was measured. These statistical metric values, which are widely used for performance evaluation, were calculated. These metrics are critical in evaluating the predictive power and accuracy of the models. The results of this study are presented in tables to compare the performance of each model and to determine which model is more effective in which situations. Table 2 presents the R^2, RMSE, MAE and MAPE values of Boosting, Bagging, RF and ANN models on the test set, allowing a comparative analysis. This table contains the necessary data to compare the performance of each model and to determine which model is more effective in which situations. These analyses will provide guidance for future studies by revealing the advantages and weaknesses of the models [67,68].

Table 7-Statistical Values by AI Models

As shown in Table-7 and Figure-3, the test data set of the models trained on the wear behavior of porcelain ceramics using ANN models with three different ensemble methods are presented to these models and the results produced by these models are visualized with graphs and statistical metrics are calculated and presented in tables. After this stage, an analysis was performed to determine which model was more successful, effective and efficient. In this analysis, the differences between the prediction results produced by each model and the actual values (residuals or error values) were calculated. These differences were plotted together on a graph including the zero line (Figure-4). This graph comparatively evaluates the overall behavior of the models and their fit to the zero line. The model with a behavior closer or like the zero line was considered to have a higher representativeness compared to the other models. This method helped to determine the most successful model by visually examining the predictive power and accuracy of the model. This evaluation process played a critical role in determining which model best predicted the wear behavior.

Figure 4: Comparison of Estimation Errors(Residuals)

When Figure-4 is analyzed, it is seen that the "Bagged" method produces worse results compared to the other two methods. The average visual achievements of the other 2 methods show similar behaviors. To make a more precise decision, statistical "Mathematical Mean" and "Standard Deviation" values were analyzed on the error values of the data produced by the models (Table-8).

Table 8: Statistical Values of Residuals

It is desirable that each error value produced here is “zero”. Therefore, the model with the arithmetic means closest to zero and the model with the lower standard deviation offer a better performance than the others. With this assumption, Figure-5 and Table-8 were prepared and presented for evaluation. This constitutes a decisive metric within these predictions.

Figure 5: Statistical Representation of Residuals

EL and ANN methodologies show significant differences in modelling and learning processes. Ensemble methods aim to reduce bias, stabilize variance and increase generalizability by combining multiple models. These methods provide more robust and precise predictions by compensating for the shortcomings of individual models. ANN is powerful in learning complex data patterns thanks to its multilayer structure and feed-forward, back-propagation learning processes. However, ANN models are often prone to overfitting and have difficulties in terms of hyperparameter adjustments and computational costs. In this study, a comparison of EL and ANN methodologies is made to contribute to the literature by evaluating the effectiveness of different modelling approaches. The differences between the ability of ensemble methods to provide overall accuracy and flexibility, and the complex data processing power of ANN are highlighted in this evaluation.

Testing Ensemble methods and ANN on the same dataset provides a reference point for future studies by demonstrating which method is more effective under specific conditions. This comparison helps to understand the differences in terms of performance criteria such as accuracy, generalization, and computational efficiency, guiding researchers in selecting the most suitable modeling approach. The study systematically evaluates the strengths and weaknesses of these methods, contributing significantly to the machine learning literature. Figure-6 presents regression plots showing the linear relationship between the values predicted by the models and the actual values. These plots visually assess the prediction accuracy and generalization ability of the machine learning models. The proximity of points to the 45-degree line indicates the prediction accuracy, while the density of points around this line reflects high accuracy. Furthermore, regression plots evaluate the model's generalization by revealing how it performs not only on training data but also on test data. Regression plots are also crucial for bias and error analysis. Significant deviations or systematic errors observed in the plots may indicate that the model is biased or overfitting specific data characteristics. In such cases, the model may need to be adjusted, or a different model should be considered. Additionally, the spread of points in the plots illustrates the error rate and distribution, providing deeper insights into the model's prediction performance.

Figure 6: Comparison of Prediction Regression Relationships

A correlation matrix is a tabular representation of the relationships between variables. It is utilized in the c domains of feature selection and engineering in machine learning, with the objective of identifying redundancies that can compromise model performance by virtue of multicollinearity. The matrix enables researchers to determine the retention, combination or removal of features, thereby enhancing the accuracy and efficiency of the model. Furthermore, it facilitates the interpretation of the model by offering insights into the data structure.

Table 9: Correlation Matrix of Models

Table-9 shows the correlation matrices for all models. This aids in interpreting artificial intelligence models that most accurately represent actual values and improves comprehension of the accompanying graphs and tables.

4. Conclusion-Discussion

In this study, the fabrication and wear properties of aluminum titanate (Al_2TiO_5) and mullite ($3Al_2O_3$-$2SiO_2$) doped porcelain ceramic-composites produced by powder metallurgy method were extensively investigated and various artificial intelligence methods were used to model these properties. The analyses include a comprehensive modeling study using ANN as well as EL methods such as Boosting, RF and Bagging. Porcelain ceramics were produced by firing at 1100-1200 ^oC for 1-5 hours using

powder metallurgy technique and their wear characteristics were evaluated through experimental testing. Steel discs served as wear discs. Wear tests were conducted on each sample with time ranging from 0-20 minutes and force between 70-120 N. The degree of wear rose with the rise in load and duration during wear testing.

Basic statistical measures such as R^2, RMSE, MAE and MAPE were used to evaluate the experimental data. The results show that Boosting performs the best among the ensemble methods, RF is close to it but performs worse, and Bagging performs the worst (Table 7). These findings make sense given the data processing logics and error correction mechanisms of each method.

Ensemble methods combine the capabilities of multiple models to enhance predictive accuracy, generalization and flexibility across a variety of data types and problem domains. Techniques such as Bagging, Boosting and RF aim to reduce the limitations of individual models and improve overall reliability by balancing bias and variance. Bagging reduces variance and increases generalizability by training independent models on subsets created by randomly resampling the original dataset, while Boosting turns weak learners into strong models by focusing on misclassified examples. Random Forests (RF) construct a multitude of decision trees with random instances and features, thereby mitigating overlearning, yet they lack the error correction mechanism present in Boosting. Boosting generally exhibits superior performance due to its capacity for error correction, while RF reduces overlearning by introducing diversity. Bagging, conversely, demonstrates a more limited effect as it concentrates solely on data sampling. These findings underscore the pivotal role of dataset structure in determining model performance and underscore the significance of selecting the most suitable ensemble method in alignment with the specific problem at hand.

When Table-7,Table-8 are analyzed, ANN have shown significant superiority compared to these three ensemble methods. ANN can learn complex data relationships and patterns more effectively due to its multilayer structure and feedforward learning algorithms. The deep learning capability of ANN provides higher accuracy and overall performance, especially on large and complex data sets. This superior performance is due to ANN's flexible architecture and learning capabilities. ANN was able to capture the complexity of the data and the relationships between features more effectively and make more successful predictions compared to Boosting, RF and Bagging.

To evaluate the performance of the models, it may not be sufficient to look only at statistical measures of value. The values of these measures summarize the data set with a single numerical value. Extreme values or possible anomalous values in the data may affect the results. For this reason, different visualizations and tables were used in the study to provide guidance to the researchers. In this section (Conclusion and discussion), numerical values were not included, instead researchers were directed to tables and figures extensively. This approach allows researchers with similar interests and questions to obtain more in-depth information.

In Figure-6, the actual values of the test data set and the values produced by the models are presented in the same graph to visually examine the overall behavior. The closeness of the ANN to the actual values here confirms the success of the statistical values. Table-7 and Table-8 are also verified with the graphs of other methods.

The values presented in Table-8 represent the arithmetic mean and standard deviation of the data groups formed by calculating the differences between the actual values and the values predicted by the models. This analysis was conducted to evaluate the alignment of the model predictions with the actual values. The key criteria here are a mean value close to zero and a low standard deviation, as small differences between the observed values and the model predictions indicate the accuracy and success of the model.

To further examine the closeness of the models to the actual values, residuals were presented using a correlation matrix (Table-9), revealing the performance of the models with the closest relationship to the actual values. This comprehensive data analysis also provides an important contribution to the literature for future research. Figure-6 illustrates the regression relationship between the predicted values of all models and the actual values, while Table-9 presents the correlation matrix values. These graphs and tables offer significant insights into the models' performance and clearly demonstrate that the best results for the dataset used in this study were achieved by the ANN.

The findings of this study contribute significantly to the development of high-performance ceramic composites and the optimization of the wear behavior of these materials. The fact that ANN emerged as the most successful model among Boosting methods highlights its ability to transform weak learners into strong models. These results provide researchers with valuable guidelines for addressing similar problems and emphasize the need for greater use of non-classical methods in future studies.

References

[1] Boyraz T.,Akkuş A.,"Investigation of wear properties of mullite and aluminum titanate added porcelain ceramics",*Journal of Ceramic Processing Research*,22[2],(2021),226-231, Doi:10.36410/jcpr.2021.22.2.226.

[2] Akkuş A.,Boyraz T.,"Fabrication and characterization of aluminum titanate and mullite added porcelain ceramics",*Journal of Ceramic Proccesing Res*.,20[1],(2019),54-58, Doi:10.36410/jcpr.2019.20.1.54

[3] Türkmen O.,Küçük A.,Akpinar S.,"Effect of wollastonite addition on sintering of hard porcelain",*Ceramics International*,41[4],(2015),5505–5512, Doi:10.1016/j.ceramint.2014.12.126.

[4] Saçlı M,Önen U,Boyraz T,"Microstructural characterization and thermal properties of aluminum titanate/porcelain ceramic matrix composites",*Acta Physica Polonica A*, 127[4],(2015),1133-1135,Doi:10.12693/aphyspola.127.1133.

[5] Kitouni S,Harabi A,"Sintering and mechanical properties of porcelains prepared from algerian raw materials",*Ceramics*,57,(2011),453-460,Doi:10.1590/S0366-69132011000400013.

[6] Martín-Márquez J.,De la Torre A.G.,Aranda M.A.G., et al.,"Evolution with temperature of crystalline and amorphous phases in porcelain stoneware",*J. Am. Ceram. Soc.*, 92,(2009),229–234,Doi:10.1111/j.1551-2916.2008.02862.x.

[7] Özsoy E,Önen U,Boyraz T,"Microstructural characterization and thermal properties of aluminum titanate/talc ceramic matrix composites",*Acta Physica Polonica A*, 127[4],(2015),1136-1138,Doi:10.12693/APhysPolA.127.1136.

[8] Çitak E,Boyraz T,"Microstructural characterization and thermal properties of aluminum titanate/ysz ceramics", *Acta Physica Polonica A*,125[2],(2014),465-468, Doi:10.12693/aphyspola.125.465.

[9] Önen U,Boyraz T,"Microstructural characterization and thermal properties of aluminum titanate/spinel ceramic matrix composites",*Acta Physica Polonica A*,125[2],(2014),488-491,Doi:10.12693/aphyspola.125.488.

[10] Martín-Márquez J.,Rincón M.,Romero M.A.G.,"Mullite development on fring in porcelain stoneware bodies", *J. Eur. Ceram. Soc.*,30,(2010),1599–1607, Doi:10.1016/j.jeurceramsoc.2010.01.002.

[11] Yahya H,Othman M.R.,Ahmad Z.A.,"Effect of mullite formation on properties of aluminosilicate ceramic balls", *Procedia Chemistry*,19,(2016),922–928, Doi:10.1016/j.proche.2016.03.136.

[12] Chargui F.,Hamidouche M.,Belhouchet H., et al.,"Mullite fabrication from natural kaolin and aluminium slag",*Boletin De La Sociedad Española De Cerámica Y Vidrio*, 57,(2018),169–177, Doi:10.1016/j.bsecv.2018.01.001.

[13] Serra M.F.,Conconi M.S.,Gauna M.R., et al.,"Mullite (3Al2O3·2SiO2) ceramics obtained by reaction sintering of rice husk ash and alumina, phase evolution, sintering and microstructure",*J. Asian Cer.Soc.*,4,(2016),61–67,Doi:10.1016/j.jascer.2015.11.003.

[14] Akkuş A.,Boyraz T.,"Investigation of wear properties of CaO, MgO added stabilized zirconia ceramics produced by different pressing methods", *J. Ceram. Proc. Res.*,19[3],(2018),249-252.

[15] Baudín C.,Tricoteaux A.,Joire H.,"Improved resistance of alumina to mild wear by aluminum titanate additions", *J. Eur. Ceram. Soc.*,34,(2014),69–80, Doi:10.1016/j.jeurceramsoc.2013.07.018.
[16] Yu H.Y.,Cai Z.B.,Ren P.D., et al.,"Friction and wear behavior of dental feldspathic porcelain",*Wear*,261,(2006),611–621,Doi:10.1016/j.wear.2006.01.024
[17] Bueno S,Micele L.,Melandri C., et al.,"Improved wear behavior of alumina–aluminium titanate laminates with low residual stresses and large grained interfaces",*J. Eur. Ceram. Soc.*,31,(2011),475–483, Doi:10.1016/j.jeurceramsoc.2010.11.002.
[18] Taktak S.,Baspinar M.S.,"Wear and friction behavior of alumina/mullite composite by sol–gel infiltration technique",*Materials and Design*,26,(2005),459–464, Doi:10.1016/j.matdes.2004.07.012
[19] Leeuwen V.,"Approaches in machine learning", *In: Verhaegh, W.F.J., Aarts, E., Korst, J. (eds) Algorithms in Ambient Intelligence. Philips Research,vol 2. Springer,Dordrecht.*, (2004),Doi:10.1007/978-94-017-0703-9_8
[20] Ryszard S.M.,Mitchell T.,"Machine learning an artificial intelligence approach",2009,*[Online]:https://api.semanticscholar.org/CorpusID:3065629*
[21] Thai H.T.,"Machine Learning for structural engineering: a state-of-the-art review", *In Structures*,38, 448-491,Elsevier,(2022),Doi:10.1016/j.istruc.2022.02.003
[22] Attou H.,Guezzaz A.,Benkirane S., et al.,"Cloud-based intrusion detection approach using machine learning techniques",*Big Data Mining and Analytics*, 6(3),311-320,(2023),Doi:10.26599/BDMA.2022.9020038
[23] Moein M.,Saradar A.,Rahmati K., et al.,"Predictive models for concrete properties using machine learning and deep learning approaches: A review",*Journal of Building Engineering*,63,105444,(2023), Doi:10.1016/j.jobe.2022.105444.
[24] Bakshi K.,"Considerations for artificial intelligence and machine learning: approaches and use cases",*IEEE Aerospace Conference,BigSky,MT:IEEE*,(2018),1-9, Doi:10.1109/AERO.2018.8396488.
[25] Aggarwal D.,Bali V.,Mittal S. ,"An insight into machine learning techniques for predictive analysis and feature selection",*Int. J. Innov. Technol. Explor. Eng.*,(2019), [Online]:https://api.semanticscholar.org/CorpusID:219632101
[26] Choudhary R.,Gianey H.K.,"Comprehensive review on supervised machine learning algorithms", *2017 Int. Conf. Mach. Learn. Data Sci. MLDS*,(2017),37-43.
[27] Mienye I.D.,Sun Y.,"A survey of ensemble learning: concepts, algorithms, applications, and prospects",*IEEE Access*,10,(2022),99129-99149, Doi:10.1109/ACCESS.2022.3207287.
[28] Sagi O., Rokach L.,"Ensemble learning: a survey", *WIREs Data Min. Knowl. Discov.*, 8,(2018),4,1249,Doi:10.1002/widm.1249.
[29] Clarke B.S,Clarke J.L.,"Predictive statistics:analysis and inference beyond models", *1.bs.Cambridge University Press*,(2018),Doi:10.1017/9781139236003.
[30] Huang F.,Xie G.,Xiao R.,"Research on ensemble learning",*International Conference on Artificial Intelligence and Computational Intelligence*, Shanghai,China:IEEE,(2009),249-252, Doi:10.1109/AICI.2009.235.
[31] Zhang Y.,Liu J.,Shen W.,"A review of ensemble learning algorithms used in remote sensing applications", *Appl. Sci.*,12[17],(2022) 8654,Doi:10.3390/app12178654.
[32] Sun Q.,Pfahringer B.,"Bagging ensemble selection", *in AI 2011:Advances in Artificial Intelligence,Ed.,Lecture Notes in Computer Science*, vol. 7106, Springer Berlin Heidelberg,(2011),251-260,Doi:10.1007/978-3-642-25832-9_26.
[33] Bühlmann P.,"Bagging, boosting and ensemble methods",*Handbook of Computational Statistics, Springer Berlin Heidelberg*,(2012),985-1022,Doi:10.1007/978-3-642-21551-3_33.

[34] Svetnik V.,Liaw A.,Tong C., et al.,"Random forest: a classification and regression tool for compound classification and QSAR modeling", *J. Chem. Inf. Comput. Sci*., 43[6], (2003),1947-1958,Doi:10.1021/ci034160g.
[35] Sanchez E.,Pintos S.,Queipo N.V.,"Toward an optimal ensemble of kernel-based approximations with engineering applications",*Struct. Multidiscip. Optim.*,36[3], (2008)247-261, Doi:10.1007/s00158-007-0159-6.
[36] Ayulani I.D.,"Tree-based ensemble methods and their applications for predicting students academic performance", *International Journal on Advanced Science, Engineering & Information Technology*,13(3)(2023), Doi:10.18517/ijaseit.13.3.16880
[37] Jung S.,Moon J.,Park S., et al.,"Bagging ensemble of multilayer perceptrons for missing electricity consumption data imputation", *Sensors (Basel, Switzerland)*,20,(2020), Doi:10.3390/s20061772.
[38] Ngo N.,Pham A.,Truong T., et al."An ensemble machine learning model for enhancing the prediction accuracy of energy consumption in buildings",*Arabian Journal for Science and Engineering*,47,4105 4117, (2021),Doi:10.1007/s13369-021-05927-7.
[39] Khan P.,Byun Y.,"Genetic algorithm based optimized feature engineering and hybrid machine learning for effective energy consumption prediction.", *IEEE Access*,8,196274-196286,(2020),Doi:10.1109/ACCESS.2020.3034101.
[40] Han H.,Zhang Z.,Cui X., et al.,"Ensemble learning with member optimization for fault diagnosis of a building energy system",*Energy and Buildings*,226,110351,(2020), Doi:10.1016/j.enbuild.2020.110351.
[41] Pinto T.,Praça I.,Vale Z., et al.,"Ensemble learning for electricity consumption forecasting in office buildings",*Neurocomputing*,423,747-755,(2021), Doi:10.1016/j.neucom.2020.02.124.
[42] Chun C.,Roy S.,Nguyen T., et al.,"Deep learning for synthetic microstructure generation in a materials-by-design framework for heterogeneous energetic materials",*Scientific Reports*,(2020),Doi:10.1038/s41598-020-70149-0.
[43] Yang S.,Meng D.,Wang H., et al.,"A novel learning function for adaptive surrogate-model-based reliability evaluation",*Philosophical Transactions of the Royal Society A*, 382(2264), (2024),Doi:org/10.1098/rsta.2022.0395
[44] Yang S.,De Jesus A.M.,Meng D., et al.,"Very high-cycle fatigue behavior of steel in hydrogen environment: state of the art review and challenges", *Engineering Failure Analysis*, 166,(2024),108898, Doi:10.1016/j.engfailanal.2024.108898
[45] Rasul A.,Karuppanan S.,Perumal V., et al.,"An artificial neural network model for determining stress concentration factors for fatigue design of tubular t-joint under compressive loads",*International Journal of Structural Integrity*,15,4,633-652, (2024), Doi:10.1108/IJSI-02-2024-0034
[46] Rasul A.,Karuppanan S.,Perumal V., et al.,"Empirical modeling of stress concentration factors using artificial neural networks for fatigue design of tubular t-joint under in-plane and out-of-plane bending moments", *International Journal of Structural Integrity*,15,4,757-776, (2024), Doi:10.1108/IJSI-03-2024-0043.
[47] Dahesh A.,Tavakkoli-Moghaddam R.,Wassan N., et al,"A hybrid machine learning model based on ensemble methods for devices fault prediction in the wood industry",*Expert Systems with Applications*, 249,123820,(2024), Doi:10.1016/j.eswa.2024.123820
[48] Freund Y.,Schapire R.E.,"A decision-theoretic generalization of on-line learning and an application to boosting",*J. Comput. Syst. Sci*.,55[1],(1997),119-39, Doi:10.1006/jcss.1997.1504.
[49] Friedman J.H.,"Greedy function approximation: a gradient boosting machine",*The Annals of Statistics*,29,5 (2001),1189-1232,https://www.jstor.org/stable/2699986

[50] Chen T.,Guestrin C., “XGBoost:a scalable tree boosting system”, *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*,(2016), [Online]:https://api.semanticscholar.org/CorpusID:4650265
[51] Zhu H.,Li T.,Zhao B.,“Statistical learning methods for neuroimaging data analysis with applications”, *Annual Review of Biomedical Data Science*,6,73-104,(2023) , Doi:10.1146/annurev-biodatasci-020722-100353.
[52] Kuncheva L.I., Whitaker C.J.,”Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy”, *Mach. Learn.*,51[2],(2003),181-207, Doi:10.1023/A:1022859003006.
[53] Manzali Y.,Elfar M.,”Random forest pruning techniques: a recent review”,*Oper. Res. Forum*, 4,43,(2023),Doi:10.1007/s43069-023-00223-6.
[54] Biau G.,Scornet E.,”A random forest guided tour”,*Test*,25[2],(2016),197-227, Doi:10.1007/s11749-016-0481-7.
[55] Sharifani K.,Amini M.,”Machine learning and deep learning: a review of methods and applications” ,*World Information Technology and Engineering Journal*,10,07,pp.3897-3904,(2023),[online] SSRN: https://ssrn.com/abstract=4458723
[56] Vlahavas I.,Vrakas D.,”*Artificial Intelligence for Advanced Problem Solving Techniques: IGI Global”*,(2008),Doi:10.4018/978-1-59904-705-8.
[57] Basheer I.A.,Hajmeer M., “Artificial neural networks: fundamentals, computing, design, and application”, *Journal of Microbiological Methods*,43[1],(2000),3-31, Doi:10.1016/S0167-7012(00)00201-3.
[58] Pantic I.,Paunovic J., Cumic J., et al.,”Artificial neural networks in contemporary toxicology research”,*Chemico-Biological Interactions*,369,(2023), Doi:10.1016/j.cbi.2022.110269.
[59] Meyers-Angulo J.E.,Martínez-Cuevas S.,Gaspar-Escribano J.M.,”Classifying buildings according to seismic vulnerability using cluster-ann techniques: application to the city of Murcia”,*Spain. Bull Earthquake Eng.*, 21,(2023) 3581–3622, Doi:10.1007/s10518-023-01671-5.
[60] Worden K.,Tsialiamanis G.,Cross E.J., et al.,”Artificial neural networks. in: rabczuk, t., bathe, kj. (eds) machine learning in modeling and simulation”,*Computational Methods in Engineering & the Sciences. Springer* ,(2023), Cham., Doi:10.1007/978-3-031-36644-4_2.
[61] Kohzadi Z.,Nickfarjam A.M.,Shokrizadeh L.A.,et al.,”A comprehensive evaluation of ensemble learning methods and decision trees for predicting trauma patient discharge status using real-world data”, *Archives of Trauma Research*,12(3),137-149,(2023), Doi:10.48307/atr.2023.407607,1020.
[62] Ahuja R.,Sharma S.C.,“Stacking and voting ensemble methods fusion to evaluate instructor performance in higher education”,*Int. J. Inf. Tecnol.*,13,1721–1731,(2021), Doi:10.1007/s41870-021-00729-4.
[63] Kulaylat A.N.,Tran L.,Kulaylat A.S., et al.,”Regression analysis”,*In Handbook for Designing and Conducting Clinical and Translational Research,Academic Press*,[26],157-163,(2023),Doi:10.1016/B978-0-323-90300-4.00087-2.
[64] Steel R.D.G.,Torrie J.H.,David A., et al.,”Principles and procedures of statistics, a biometrical approach”,*3rd Edition,McGraw Hill, Inc. Book Co.*,New York,(1997),352-358.
[65] Chicco D.,Warrens M.J.,Jurman G.,et al.,”The coefficient of determination r-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation”.*PeerJ Computer Science*,7,623,(2021), Doi:10.7717/peerj-cs.623.

[66] Chai T., Draxler R.R.,"Root mean square error (RMSE) or mean absolute error (MAE) – arguments against avoiding RMSE in the literature",*Geosci. Model Dev.*,7[3],(2014),1247-1250, Doi:10.5194/gmd-7-1247-2014.
[67] Ganaie M.A.,Hu M.,Malik A.K., et al.,"Ensemble deep learning: a review",*Eng. Appl. Artif. Intell.*,115,(2022),105151, Doi:10.1016/j.engappai.2022.105151.
[68] Mohammed A., Kora R.,"A comprehensive review on ensemble deep learning:opportunities and challenges",*J. King Saud Univ.,Comput. Inf. Sci.*,35[2],(2023),757-774, Doi:10.1016/j.jksuci.2023.01.014.

Assoc.Dr. Ahmet Gürkan YÜKSEK Biography

Assoc. Prof. Dr. Ahmet Gürkan Yüksek is a distinguished academic specializing in artificial intelligence, the Internet of Things (IoT), and embedded systems. He serves as a faculty member at the Department of Computer Engineering, Faculty of Engineering, Sivas Cumhuriyet University. Additionally, he holds the position of Director at the Artificial Intelligence and Data Science Application and Research Center. Throughout his academic career, Dr. Yüksek has developed solutions to various engineering problems using artificial neural networks, fuzzy logic, and machine learning methods. His research interests extend to materials science and engineering, with notable studies on predicting the wear properties of zirconia ceramics and modeling the temperature-dependent photoluminescence properties of GaN epilayers.

Dr. Yüksek has published numerous articles in journals indexed by SCI-Expanded, contributing significantly to the advancement of artificial intelligence applications in engineering and technology. He continues to serve the scientific community through his academic and administrative roles, fostering innovation and research in his fields of expertise.

Figure and Table captions

Figure and Tables

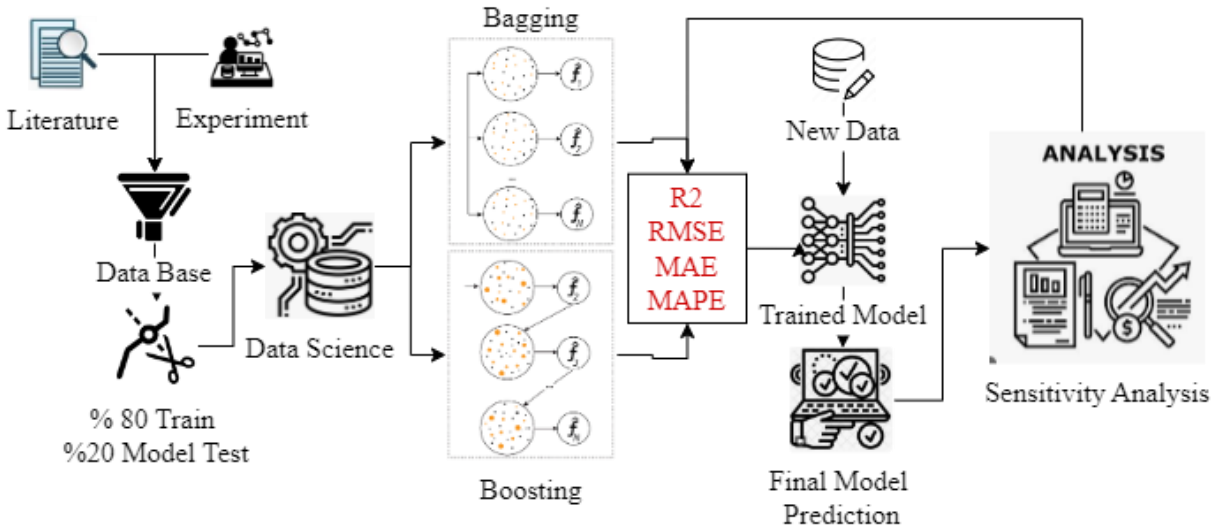


Figure 1: General Process Design of the Proposed Model

Table 1: Data Set Parametric Structure

					Wear Volume (10^{-2},mm^3)			
	Force(N)	**Wear Time(min)**	**Temp.**	**Temp.Time**	**P**	**PM**	**PAT**	**PMAT**
Min-Max	70-120	5-20	1100-1200	1- 5	3.00-33.33	0.53-7.41	0.32-5.48	1.06-38.44
Average	93.3	12.5	1150	3	12.6	3.03	1.9	14
Std.Dev.	20.6	5.62	41	1.64	7.3	1.8	1.2	10
	$Input_1$	$Input_2$	$Input_3$	$Input_4$	Models Output			

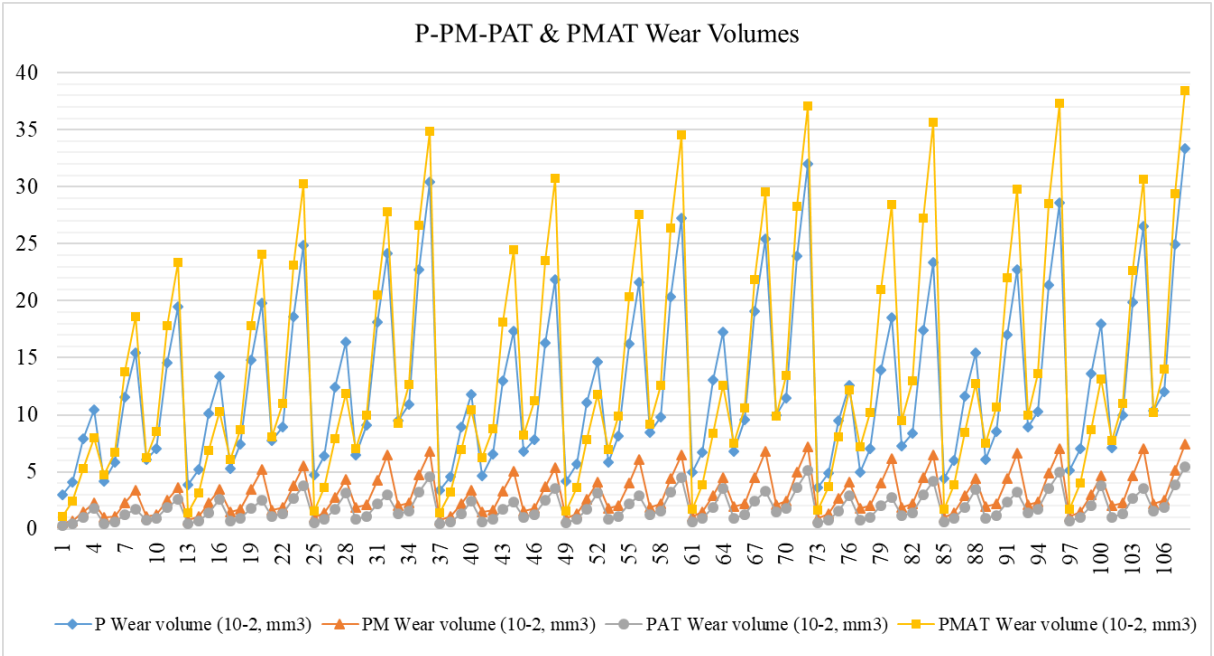


Figure 2: Wear Behavior of Porcelain-Ceramic Specimens

Table 2: The Detail of Boosting

Initialize Weights, $\{w_i\}_{i=1}^{n}$	w_i=1/N for i=1,..,N
For each iteration t:	• Train a weak learner $h(x)$ on the weighted training set. • Calculate the error et of $h_t(x)$: $e_t = \frac{\sum_{i=1}^{n} w_i . I(y_i \neq h_t(x_i))}{\sum_{i=1}^{n} w_i}$ Equation5 • Compute the learner's weight α_t: $\propto_t = \ln(\frac{1-e_t}{e_t})$ Equation6 • Update the weights for the next iteration: $w_i \longleftarrow w_i . \exp(\propto_t . I(y_i \neq h_t(x_i)))$ Equation7 Normalize the weights so that they sum to 1. • The final model $H(x)$ is a weighted sum of the weak learners: $H(x) = \sum_{t=1}^{T} \propto_t . h_t(x)$ Equation8

Table 3: The Detail of Bagging

Bootstrap Sample Generation:	Original dataset $D = \{(x_i, y_i)\}_{ib=1}^{n}$ Equation9
For b=1 to B:	• Generate $D_b = \{(x_i, y_i)\}_{ib=1}^{n}$, where x_{ib} drawn from D with replacement • Note that some instances from D may appear multiple times in , while others may not appear at all.
Model Training	• For each bootstrap sample Db: o Train a model Mb on Db. o The model Mb is fitted to minimize the error on Db.
Prediction Aggregation	• For a new instance x, the ensemble prediction is the average of the predictions from all models Mb: $y' = \frac{1}{B}\sum_{b}^{B} M_b(x)$ Equation10

Table 4: The Detail of RF

Bootstrap Sampling:	Generate multiple bootstrap samples from the original dataset by randomly sampling with replacement. Given a training set $D = \{(x_i, y_i)\}_{i=1}^{n}$, create B bootstrap samples D_b
Random Feature Selection	For each tree, select a random subset of features at each split. If there are p features, typically $\sqrt{p}$ features are chosen for classification, and $p/3$ for regression.
Training Multiple Trees	Train a separate decision tree Tb on each bootstrap sample Db using the randomly selected subset of features.
Aggregate Prediction	Combine the predictions of all trees to make the final prediction. $y' = \frac{1}{B}\sum_{b=1}^{B} T_b(x)$ Equation11

Table 5: The Detail of ANN

Forward Propagation	$z^{(l+1)} = w^{(l)}a^{(l)} + b^{(l)}$ $a^{(l+1)} = f(z^{(l+1)})$ Equation12
Back Propagation	$\delta^{(l)} = \nabla aC \odot f'(z^{(l)})$ $\delta^{(l)} = ((w^{(l)})^T \delta^{(l+1)}) \odot f'(z^{(l)})$ $\frac{\partial w^{(l)}}{\partial C} = \delta^{(l+1)}(a^{(l)})^T$ $\frac{\partial b^{(l)}}{\partial C} = \delta^{(l+1)}$ Equation13

Table 6: Pseudo Code for Artificial Intelligence Models [62]

Initialize weights for all instances in the training set For each iteration: Train a weak learner on the weighted training set Evaluate the weak learner's performance Increase weights of misclassified instances Decrease weights of correctly classified instances Combine weak learners to form a strong learner a- Boosting	For each of N models: Generate a bootstrap sample of the training data Train the model on the bootstrap sample For each instance in the test set: Aggregate predictions from all models (average for regression) b- Bagging
For each of N trees: Generate a bootstrap sample of the training data Select a random subset of features for each split in the tree Train the decision tree on the bootstrap sample For each instance in the test set: Aggregate predictions from all trees (average for regression) c- Random Forest	Initialize the neural network parameters (weights and biases) For each epoch: For each batch of training data: Perform forward propagation to compute predictions Compute the loss between predictions and actual values Perform backward propagation to compute gradients Update parameters using gradients Evaluate the model on the test set d- ANN

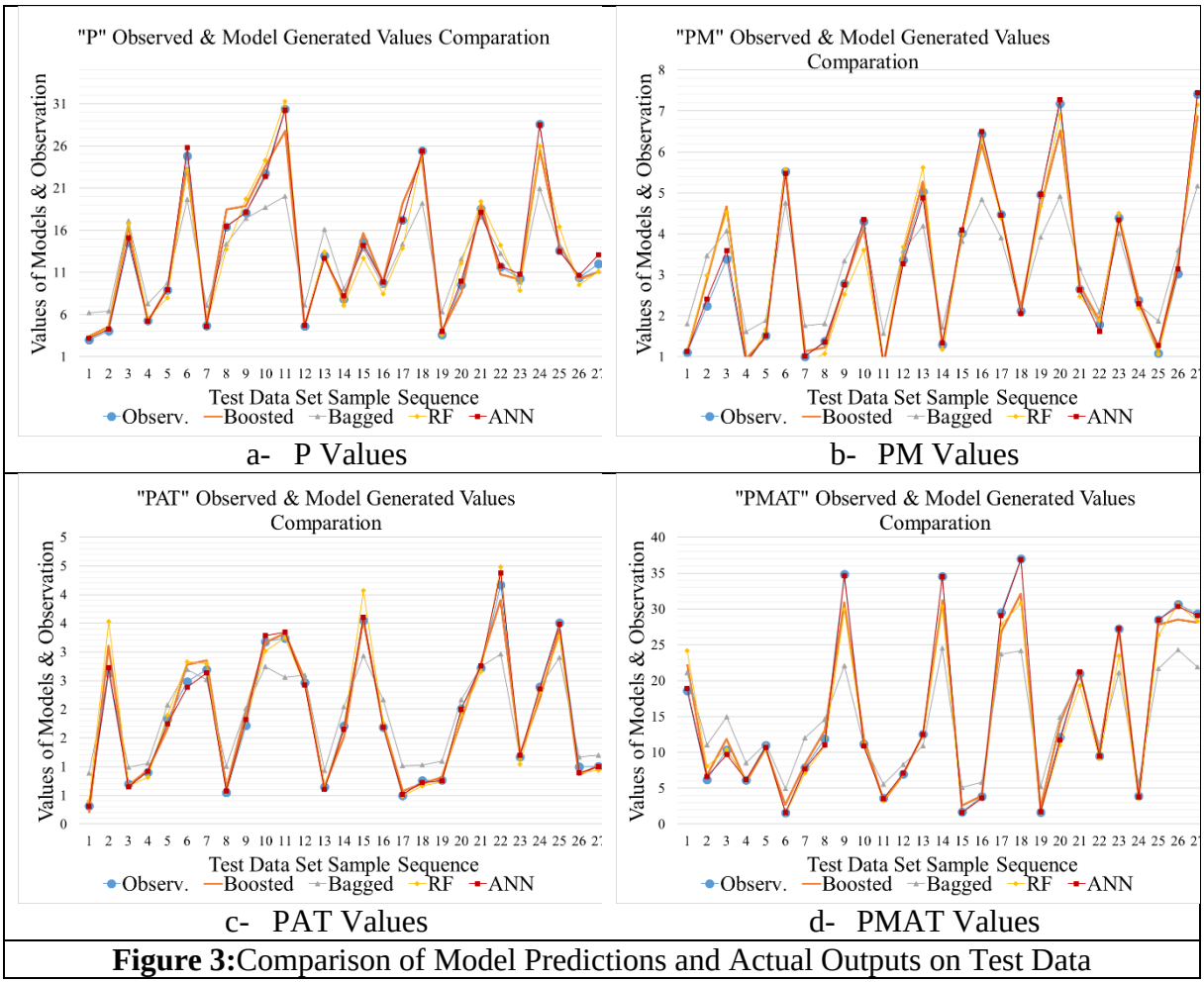
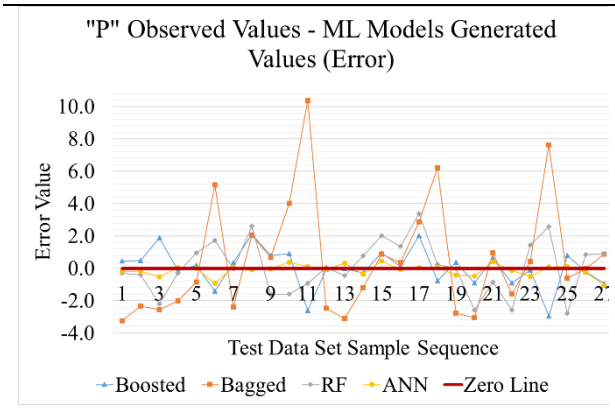


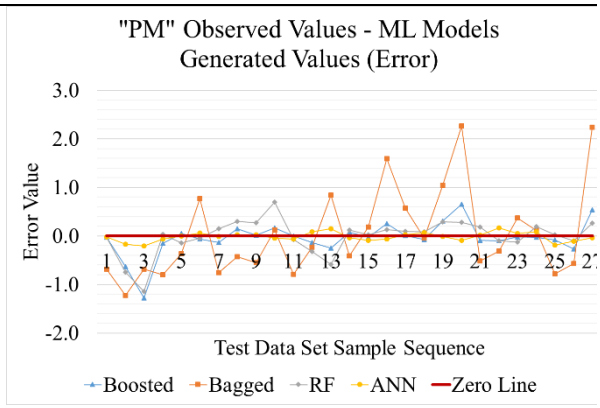
Figure 3: Comparison of Model Predictions and Actual Outputs on Test Data

Table 7: Statistical Values by AI Models

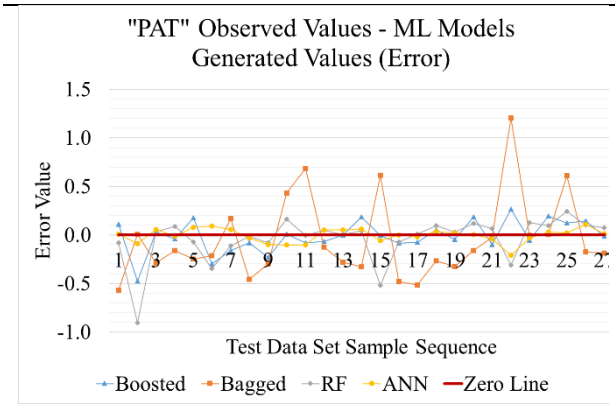
	Indicator	Boosted Ensemble		Bagged Ensemble		Random Forest		Artificial Neural Network	
P	R^2	0.98488	0.97644	0.83631	0.78794	0.96477	0.95406	0.99908	0.99749
	RMSE	0.87478	1.16970	2.87848	3.50906	1.33547	1.63328	0.20256	0.38158
	MAE	0.52077	0.86389	2.03866	2.61573	0.98497	1.31327	0.16233	0.27357
	MAPE	4.02750	6.38740	19.82040	25.27710	8.12863	9.95413	1.88410	2.75690
PM	R^2	0.98245	0.96849	0.80865	0.78234	0.94102	0.96637	0.99967	0.99773
	RMSE	0.23292	0.34361	0.76913	0.90307	0.42702	0.35495	0.02904	0.09229
	MAE	0.15193	0.20662	0.56680	0.71388	0.24931	0.24240	0.02354	0.07310
	MAPE	6.08680	7.15020	24.69850	30.64980	9.43447	8.75949	1.02420	3.35960
PAT	R^2	0.98088	0.97785	0.82156	0.85197	0.96372	0.95238	0.99928	0.99574
	RMSE	0.17029	0.16131	0.52017	0.41702	0.23454	0.23652	0.03384	0.07072
	MAE	0.11466	0.12205	0.36891	0.32749	0.17091	0.14306	0.02572	0.05478
	MAPE	6.52230	8.08830	25.23130	29.13860	10.86685	7.65816	1.54270	3.40240
PMAT	R^2	0.98395	0.97536	0.81566	0.78178	0.96185	0.95961	1.00000	0.99937
	RMSE	1.18190	1.81071	4.00576	5.38865	1.82223	2.31822	0.00237	0.29049
	MAE	0.76656	1.25806	3.11435	4.20407	1.18336	1.43725	0.00173	0.21968
	MAPE	9.26950	12.33110	44.07530	49.12520	9.82950	8.31539	0.03464	2.36500
		Train	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**



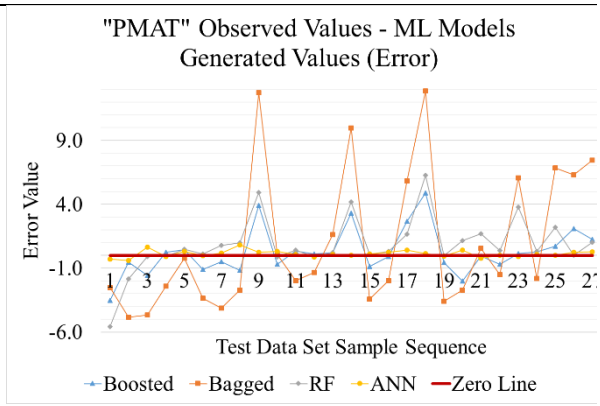
a- P Values



b- PM Values



c- PAT Values

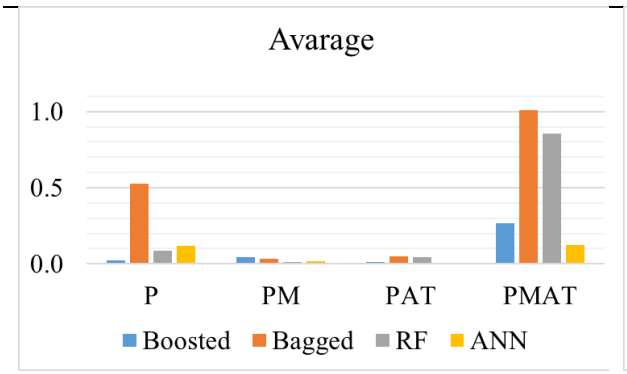


d- PMAT Values

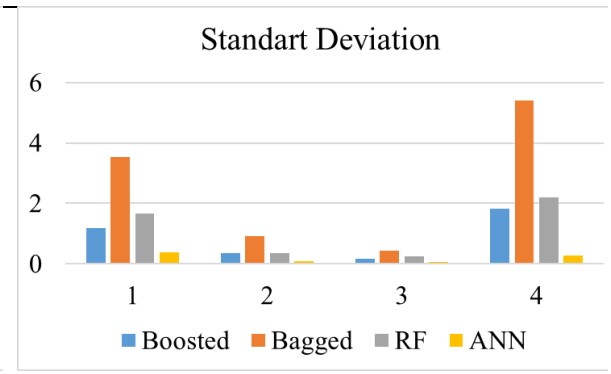
Figure 4: Comparison of Estimation Errors(Residuals)

Table 8: Statistical Values of Residuals

	Average				Standard Deviation			
	P	PM	PAT	PMAT	P	PM	PAT	PMAT
Boosted	0.0236	0.0421	0.0132	0.2652	1.1917	0.3475	0.1638	1.8253
Bagged	0.5284	0.0350	0.0514	1.0101	3.5351	0.9196	0.4217	5.3940
RF	0.0850	0.0103	0.0437	0.8549	1.6621	0.3616	0.2369	2.1959
ANN	0.1207	0.0158	0.0046	0.1232	0.3689	0.0927	0.0719	0.2681

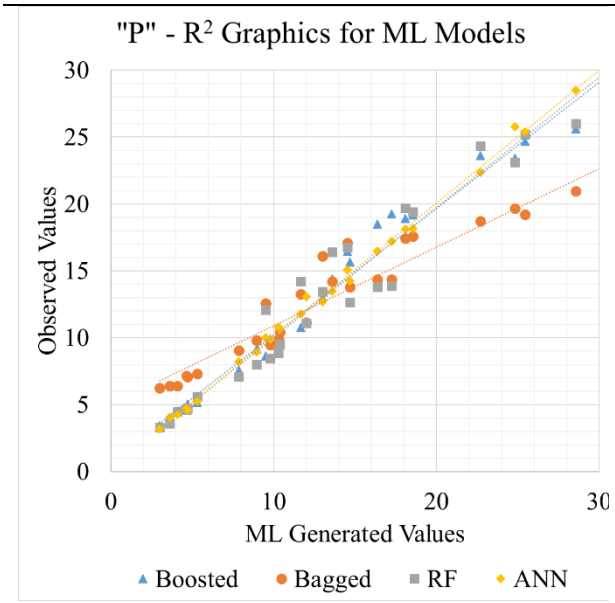


a- Average

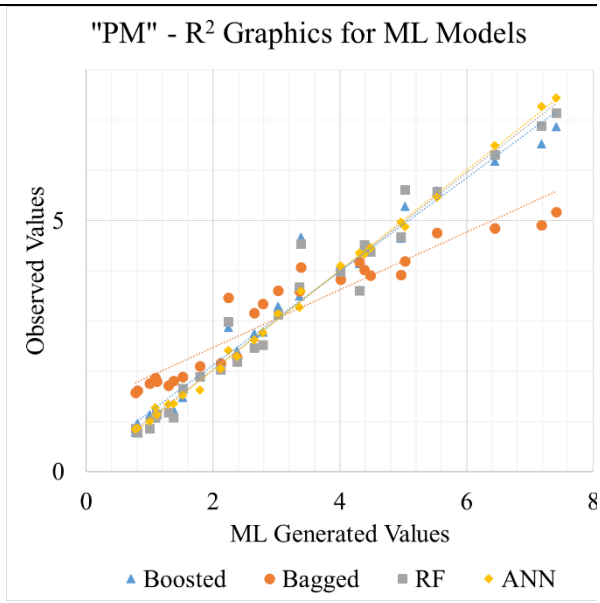


b- Standard Deviation

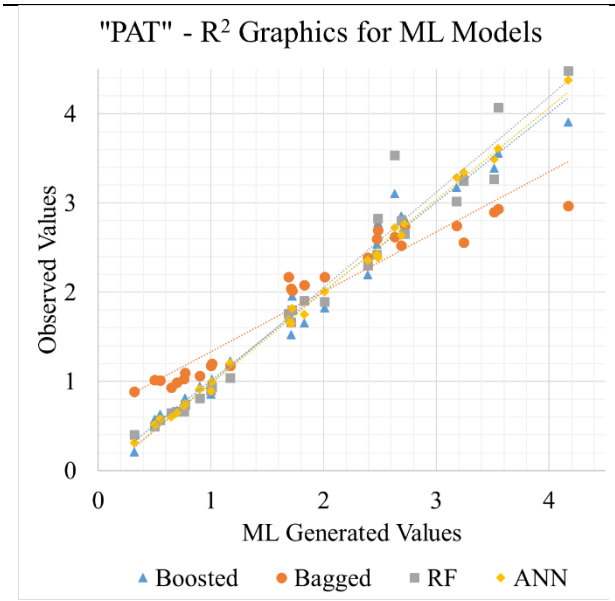
Figure 5: Statistical Representation of Residuals



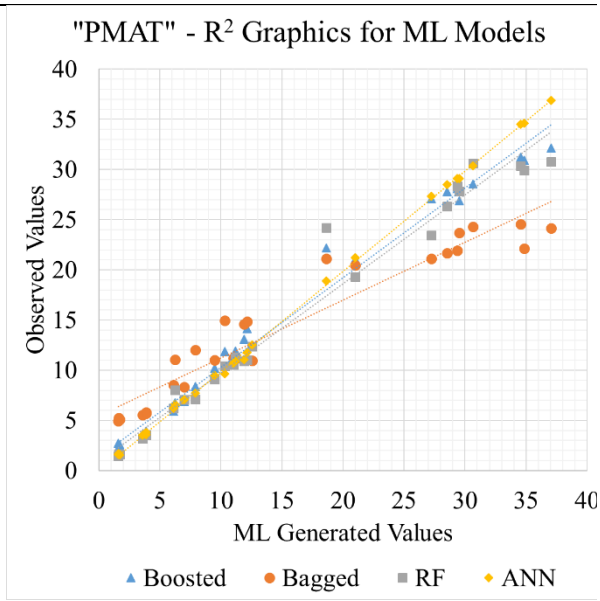
a- P Values



b- PM Values



c- PAT Values



d- PMAT Values

Figure 6: Comparison of Prediction Regression Relationships

Table 9: Correlation Matrix of Models

P	*Observed*	*Boosted*	*Bagged*	*RF*	*ANN*
observed	1.000				
Boosted	0.988	1.000			
Bagged	0.951	0.957	1.000		
RF	0.977	0.964	0.969	1.000	
ANN	0.998	0.985	0.949	0.975	1.000

PM	*observed*	*Boosted*	*Bagged*	*RF*	*ANN*
observed	1.000				
Boosted	0.985	1.000			
Bagged	0.948	0.975	1.000		
RF	0.983	0.995	0.962	1.000	
ANN	0.998	0.986	0.952	0.983	1.000

PAT	*observed*	*Boosted*	*Bagged*	*RF*	*ANN*
observed	1.000				
Boosted	0.989	1.000			
Bagged	0.960	0.964	1.000		
RF	0.982	0.987	0.952	1.000	
ANN	0.998	0.988	0.952	0.983	1.000

PMAT	*observed*	*Boosted*	*Bagged*	*RF*	*ANN*
observed	1.000				
Boosted	0.992	1.000			
Bagged	0.957	0.977	1.000		
RF	0.986	0.991	0.973	1.000	
ANN	0.999	0.992	0.956	0.986	1.000