

Article

Evaluation of the Effect of Using Different Types of Clinker Grinding Aids on Grinding Performance by Numerical Analysis

Yahya Kaya ¹, Veysel Kobya ¹, Murat Eser ² , Naz Mardani ³, Metin Bilgin ⁴  and Ali Mardani ^{1,*} 

¹ Department of Civil Engineering, Bursa Uludag University, Bursa 16059, Turkey; yahyakaya00@gmail.com (Y.K.); v.koby@gmail.com (V.K.)

² Department of Computer Engineering, Canakkale Onsekiz Mart University, Canakkale 17010, Turkey; meser@comu.edu.tr

³ Department of Mathematics Education, Bursa Uludag University, Bursa 16059, Turkey; nazmardani@uludag.edu.tr

⁴ Department of Computer Engineering, Bursa Uludag University, Bursa 16059, Turkey; metinbilgin@uludag.edu.tr

* Correspondence: alimardani@uludag.edu.tr

Abstract: To develop more environmentally friendly and sustainable cementitious systems, the use of grinding aids (GAs) during the clinker grinding process has increasingly gained attention. Although the mechanisms of the action of grinding aids (GAs) are known, the selection of an effective grinding aid (GA) can be difficult due to the complexity of appropriate selection criteria. For this reason, it is important to model the effect of GA properties on grinding performance. In this study, seven different types of GAs were used in four different dosages, and time-dependent grinding was performed. The Blaine fineness values of cements were compared after each grinding process. In addition, the modeling of these parameters using machine learning and ensemble learning methods was discussed. The Synthetic Minority Over-sampling Technique (Smote) was used to generate artificial data and increase the number of data for the grinding efficiency experiment. The data were modeled using methods such as Artificial Neural Networks (ANNs), Attentive Interpretable Tabular Learning (TabNet), Random Forests (RFs), and the XGBoost Regressor (XGBoost), and the ranking of the parameters affecting the Blaine properties was determined using the XGBoost method. The XGBoost method achieved the best results in the MAE, RMSE, and LogCosh metrics with values of 21.0384, 33.7379, and 15.4846, respectively, in the experimental modeling studies with augmented data. This study contributes to a better understanding of the relationship between GA selection and milling process performance.

Keywords: grinding efficiency; grinding aids; Blaine fineness; artificial neural networks; TabNet; XGBoost; Random Forest; machine learning; regression analysis



Academic Editors: Predrag Dašić, Adolfo Senatore, Alessandro Ruggiero and Geo Paul

Received: 13 May 2025

Revised: 1 June 2025

Accepted: 4 June 2025

Published: 9 June 2025

Citation: Kaya, Y.; Kobya, V.; Eser, M.; Mardani, N.; Bilgin, M.; Mardani, A. Evaluation of the Effect of Using Different Types of Clinker Grinding Aids on Grinding Performance by Numerical Analysis. *Materials* **2025**, *18*, 2712. <https://doi.org/10.3390/ma18122712>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The growing demand for cement contributes significantly to environmental issues, primarily due to CO₂ emissions generated during its production. Various strategies have been adopted to mitigate the adverse environmental impacts associated with cement use. One such approach involves enhancing the compatibility between cement and water-reducing admixtures to enable more sustainable and eco-friendly production processes [1–6]. Another strategy focuses on the partial or complete replacement of cement with alternative binders [7–11]. A third method is the incorporation of grinding aids (GAs) during the

cement grinding process, which can enhance both grinding efficiency and the performance characteristics of cement [12–14].

GAs are used to save energy and reduce greenhouse gas emissions during clinker grinding [15,16]. GAs are adsorbed on clinker particle surfaces due to their high-polarity functional groups ($-\text{OH}$, $-\text{NH}_2$, $-\text{COOR}$, $-\text{SO}_3$, etc.) [17,18]. An adsorbed GA neutralizes the electrical charges on the surface and prevents clinker cracks from closing and particles from coalescing and sticking to the mill surface and/or balls [17,19]. This saves energy during the grinding stage. In clinker grinding, amine and glycol-based GAs are generally used to save energy [20–23]. Along with the development of this technology, various modification studies have been carried out to improve the performance of GAs. For instance, some studies found that the hydroxyl groups of GAs are enhanced by various modification processes to improve their performance [24,25].

Although the mechanisms of action of GAs are known, the selection of these additives is still based on hypothetical information. In particular, due to the complexity of the appropriate selection criteria, the selection of a performance-efficient grinding aid is extremely difficult. Many parameters such as molecular weight, number of hydrogen bonds, pH, number of functional groups, etc., can affect grinding performance. For this reason, modeling the effect of GA properties on grinding performance becomes important in terms of both time and cost. Modeling these parameters can be achieved using two different learning methods, machine learning and ensemble learning. Machine learning is a field of study that focuses on developing algorithms and models that allow computers to learn and make predictions or decisions without being explicitly programmed. It is also used in natural language processing, computer vision, regression problems, fraud detection, recommender systems, and many other areas [26–31]. It performs statistical computations on data to detect, learn, and predict specific patterns in a dataset. Ensemble learning is an approach to machine learning that makes predictions based on the decisions of multiple learning algorithms or learning methods that come together to solve a problem. The ensemble methods are trained separately to solve the problem. In the decision phase, a final decision is made based on the outputs produced by the methods [32,33]. The main purpose of its use is to try to increase the prediction success of the method. Some modeling studies on GAs and the grinding process are given in Table 1.

Table 1. Some modeling studies on GAs and the grinding process.

Reference	Estimated Parameter	Methodology Used	Highlights
[34]	Modeling and optimization of PEA-combined grinding aids' effect in the grinding circuit of cement ball mills	Quadratic model for regression analysis	A quadratic regression model was used to investigate the effects of key parameters such as mill utilization rate, Blaine, and GA dosage rate to find the most efficient conditions.
[35]	Grinding roughness prediction model based on an evolutionary artificial neural network.	Genetic algorithm and artificial neural network	Based on analyzing the Back Propagation (BP) network disadvantages of low convergence speed and frequently falling into local minimum value, a roughness prediction model of a BP neural network integrated with a genetic algorithm was proposed.

Table 1. *Cont.*

Reference	Estimated Parameter	Methodology Used	Highlights
[36]	Application of statistical and soft computing techniques for the prediction of grinding performance	Artificial neural networks and regression models	In this paper, statistical methods and soft computing techniques, namely regression models completed with analysis of variance and artificial neural networks, respectively, are presented for the estimation of grinding forces and temperature.
[37]	Identification and expert approach to controlling the cement grinding process using artificial neural networks and other non-linear models	NARX models based on feed-forward networks, Elman, Jordan, and Layered Recurrent Network (LRN) recurrent networks, as well as MTL (Multi-Task Learning) and traditional NARX non-linear models	The paper involved conducting preliminary research to explore the identification and control of a multi-dimensional, non-linear, and non-stationary cement grinding process using artificial neural networks and various other non-linear models.
[38]	Empirical mode decomposition-based hybrid ensemble model for electrical energy consumption forecasting of the cement grinding process	Empirical mode decomposition (EMD), moving average filter (MAF), least squares support vector regression (LSSVR), and quadratic exponential smoothing (QES)	Forecasting the electrical energy consumption of the cement grinding process remains a difficult task due to the intrinsic complexity and irregularity of its time series. To solve this difficulty and improve the prediction accuracy, a novel hybrid model was proposed based on the “decomposition-prediction-integration” methodology.

This study addresses a critical challenge in modeling experimental data for grinding efficiency: the limited size of the original dataset. To overcome this, the Synthetic Minority Over-sampling Technique (SMOTE) was employed to generate high-quality artificial data, effectively enriching the dataset and enhancing model robustness. Unlike simple data duplication, SMOTE innovatively creates new synthetic samples by interpolating between existing minority class data points using K-Nearest Neighbors, thus preserving the underlying data distribution and improving learning outcomes [39]. Leveraging this augmented dataset, grinding efficiency was modeled using a combination of advanced learning techniques, including Artificial Neural Networks (ANNs) as a traditional machine learning method, the transformer-based Attentive Interpretable Tabular Learning (TabNet) for deep learning, and ensemble methods such as Random Forest (RF) and the XGBoost Regressor. This multi-method approach not only improves prediction accuracy but also provides comprehensive insights into the factors controlling Blaine fineness, with the Random Forest model specifically employed to rank the dominance of input parameters influencing grinding performance. The integration of SMOTE with diverse state-of-the-art modeling techniques in this work represents a novel and systematic strategy to optimize data-driven predictions in grinding aid research, contributing valuable methodological advancements to the field.

2. Materials and Methods

2.1. Materials

Cements containing different types and proportions of GA were produced as part of this study. For this purpose, 96% clinker and 4% gypsum were used. A total of 85 cements with a GA were prepared following the EN 197-1 standard [40] by using different GAs during grinding processes carried out at three different grinding times in a laboratory-type ball mill. Accordingly, in addition to the commonly used commercial GAs based on triethanolamine (TEA), triisopropanolamine (TIPA), diethanolisopraponalamine (DEIPA), diethylene glycol (DEG), and ethylene glycol (EG), a total of 7 different GAs, including two different modified TEAs obtained by the esterification reaction of a TEA-type GA, were added to the mixture during the cement grinding stage. All the GAs were used in proportions of 0.025%, 0.05%, 0.075%, and 0.1% of the total clinker and gypsum mass. The properties of the GAs used are shown in Table 2, and the Blaine finenesses of each cement after grinding at three different grinding times are shown in Figure 1. The cements were named according to the type and dosage of the GA used. For example, the cement prepared with 0.025% TEA was named 0.025 TEA.

Table 2. Some chemical properties of the GAs used.

Types of GAs	Molecular Weight (g/mol)	Density (g/cm ³)	pH. 25 °C	Hydrogen Bonds	Number of Functional Groups
TEA	148.19	1.095	10.5	4	3
TIPA	191.27	1.124	10.8	4	3
DEIPA	163.20	1.079	9.7	3	3
DEG	106.12	1.118	7.2	3	2
EG	62.07	1.26	8.2	3	2
M-TEA1	207.00	1.206	5.6	4	3
M-TEA2	327.00	1.166	2.63	4	3

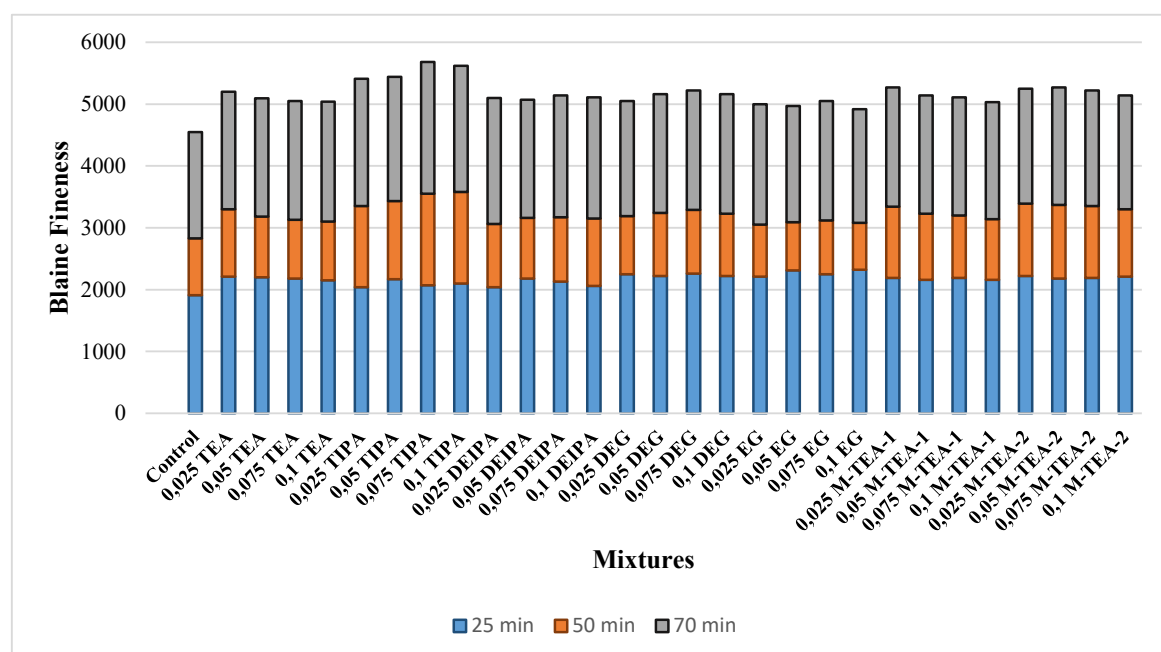


Figure 1. Blaine fineness values of cement prepared with different grinding times, GA type, and dosage.

2.2. Method

Cement production was carried out using a laboratory-type ball mill with 1.5 kW motor power, 35 rpm rotation speed, 19 cm width, 48 cm diameter, and 5 kg capacity (Figure 2). Clinker and gypsum were prepared according to the Bond Standard before grinding. Ball distribution was also selected according to the Bond Standard. In addition, the Blaine fineness of the ground cements was measured using an automatic Blaine apparatus, depicted in Figure 3.

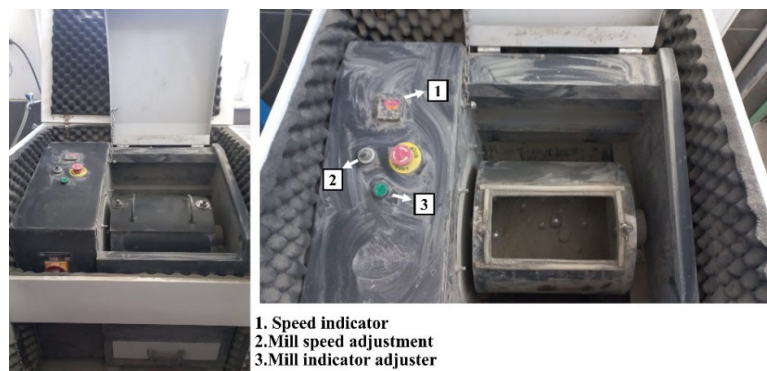


Figure 2. Laboratory type mill where the grinding process is carried out.



Figure 3. Automatic Blaine device.

Methodology of Modeling

Information on the modeling of Blaine data obtained as a result of experimental studies using 8 different methods in a computer environment is presented. The 84 data obtained from the experimental studies were used in the modeling studies. To create more efficient models, the input data were first normalized between 0.1 and 0.9. The corresponding formula is shown in Equation (1). Since machine learning algorithms are sensitive to the differences between the scales of the input data, learning with unscaled data usually results in poor performance. In multivariate regression problems, normalization is used to improve performance so that variables in different scale ranges contribute equally to the analysis. For example, if normalization is not applied to a feature between 0 and 1000 and a feature between 0 and 10 in the dataset, the distribution of weights for the model will be disproportionate. For this reason, the goal is to increase the success rate by applying the normalization process to experimental data.

$$O_i = 0.8 \frac{(O - O_{min})}{(O_{max} - O_{min})} + 1 + 0.1 \quad (1)$$

Before computer modeling of the Blaine fineness values of the cements obtained after grinding, artificial data were generated by the Smote method using the available data. After the data augmentation process, the number of data for Blaine experiments was increased from 84 to 1500. For the determination of the hyperparameters of the XGBoost and RF models, a K-fold cross-validation technique was used to ensure that all data were available in both the training and test sets. K-fold cross-validation is a widely used technique for estimating the performance of a machine learning algorithm on a given dataset. It involves dividing the dataset into K subsets, using K-1 subsets for training, and the remaining subset for validation. This process is repeated K times, and each subset is used once as validation data. The results are then averaged to produce a single prediction [41]. The purpose of using cross-validation in machine learning is to accurately predict the performance of machine learning algorithms and evaluate their generalization capabilities. Cross-validation techniques, such as K-fold cross-validation, are essential for evaluating the predictive performance of machine learning models [36–38]. It is particularly important because it provides a principled framework for selecting models that maximize generalization performance [20,21]. Cross-validation is used to avoid overfitting and to ensure that model performance is not overly optimistic or pessimistic [41–43].

In summary, while using machine learning for modeling, data preprocessing (normalization and augmentation with SMOTE) [44–46], modeling, model evaluation (K-fold cross-validation) steps [47–49] were systematically applied on Blaine fineness data. All operations performed in the modeling step were performed to ensure model stability and reliability. Python programming language version 3.9, sklearn library version 1.6.1, XGBoost library version 3.0.1 and imbalanced-learn library version 0.13 were used in the modeling phase.

In this study ANN, TabNet, RF, and XGBoost methods were used as modeling methods. Equation (2) was used to determine the coefficients in the Linear Regression (LR) method [50]. In Equation (2), M: Number of samples; p: features; w: coefficients; y_i : actual value; $\underline{y}_i$: predicted value; and x: inputs to adjust the coefficients. In determining the parameters of the modeling methods, the parameters that each method can take were prepared and the most effective parameters of the relevant method on the dataset were selected by trying all parameter combinations on the dataset. The ranking of influential parameters plays a crucial role in both selecting the appropriate GA for the grinding process and enhancing grinding efficiency. Accurate prioritization of these parameters facilitates more environmentally sustainable production under economically favorable conditions, as focusing on the most dominant factors can lead to significant savings in time, energy, and cost. While adjusting the hyperparameters in the ANN method, the number of hidden layers is between 2 and 5, the optimization algorithm is Adam, and SGD and Relu and Linear are used as activation functions. In TabNet method, the learning coefficient was used between 2×10^{-1} – 2×10^{-3} , step size was between 5 and 20, gamma was between 0.5 and 1.0, and the number of epochs used was 500. In the XGBoost method, when deciding the number of predictors between 3 and 49, squared_error, absolute_error, friedman_mse, and Poisson maximum number of depths was 1–5 for the criterion method used and 0.1–1.0 was for the learning coefficient, hist, exact, and approx metrics for the maximum tree method parameter. In the RF method, when deciding the number of estimators between 3 and 99, the metrics squared_error, absolute_error, friedman_mse, and Poisson maximum depth of 1–5 were used for the criterion method.

$$\sum_{i=1}^M \left(y_i - \underline{y}_i \right)^2 = \sum_{i=1}^M \left(y_i - \sum_{j=0}^p w_j \times x_{ij} \right)^2 \quad (2)$$

Three different metrics were used to evaluate the results of 3 different modeling methods applied within the scope of this study. Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Logcosh metrics shown in Equations (3)–(5) were used to evaluate the modeling results.

$$MAE = \frac{1}{n} \sum_{i=1}^n |E_i - P_i| \quad (3)$$

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (E_i - P_i)^2 \right]^{\frac{1}{2}} \quad (4)$$

$$Logcosh = \sum_{i=1}^n \log(\cosh((E_i - P_i))) \quad \cosh(x) = \frac{e^x + e^{-x}}{2} \quad (5)$$

Here, E_i is the experimental result, P_i is the result obtained from our model, and n is the number of data.

The flowchart of the study is given in Figure 4.

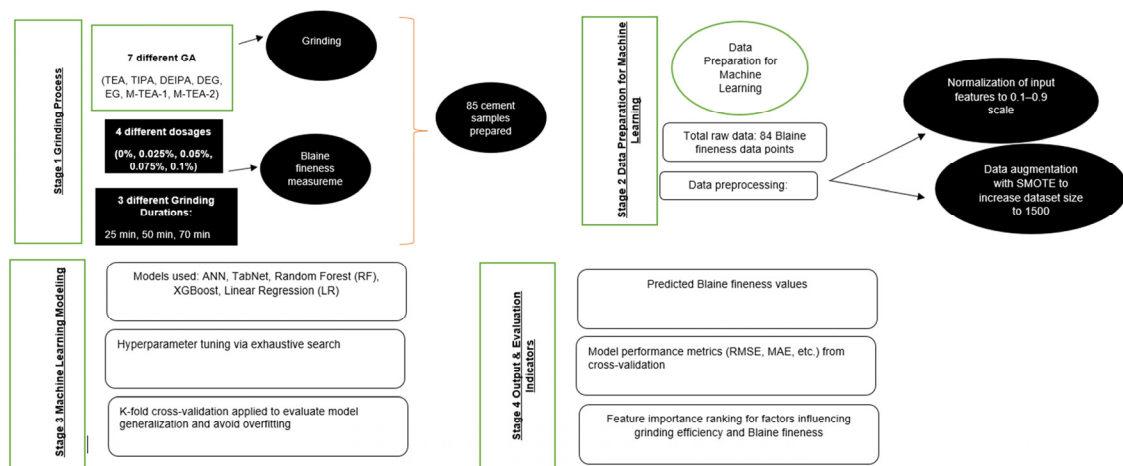


Figure 4. Flowchart of this study.

3. Result and Discussion

3.1. Assessment of Experimental Results

As illustrated in Figure 1, the Blaine fineness of all cement samples increased with extended grinding time from 25 to 50 min, independent of GA usage. After 25 min of grinding, the control cement (without GA) exhibited the lowest Blaine fineness, whereas the incorporation of GAs led to significant enhancements. Specifically, compared to the control, the Blaine fineness increased by 14%, 9%, 11%, 17%, 19%, 14%, and 16% for TEA, TIPA, DEIPA, DEG, EG, M-TEA-1, and M-TEA-2, respectively. These improvements confirm the efficiency of GA use in short-term grinding, particularly for glycol-based EG, which outperformed others, while TIPA demonstrated the lowest performance in this phase. This trend aligns with findings by Assaad et al. [51], who reported superior short-term grinding efficiency for glycol-based GAs.

Regarding dosage sensitivity, no significant improvement in grinding efficiency was observed with increasing GA dosage beyond optimal levels. For amine-based GAs (TEA, TIPA, DEIPA, M-TEA-1, M-TEA-2), the highest Blaine values were obtained at dosages of 0.025–0.05%, whereas glycol-based GAs (DEG, EG) achieved maximum efficiency at 0.075% and 0.1%, respectively. This suggests that amine-based GAs can reach optimal performance at lower dosages, offering a more dosage-efficient grinding process in the short term.

For the 50 min grinding condition, the control sample again showed the lowest fineness. In contrast, the use of TEA, TIPA, DEIPA, DEG, EG, M-TEA-1, and M-TEA-2 led to fineness increases of 14%, 22%, 10%, 14%, 9%, 14%, and 18%, respectively. Notably, EG, which was the most effective GA at 25 min, became the least efficient at 50 min, whereas TIPA emerged as the most effective. This performance shift is attributed to the ability of TIPA to reduce surface energy through its hydroxyl groups, as supported by research findings [52].

Grinding efficiency was also dosage-dependent in this phase. TIPA exhibited improved performance with increasing dosage, while DEIPA, DEG, and EG reached optimal effectiveness at 0.075%. For TEA, M-TEA-1, and M-TEA-2, an inverse relationship was observed; higher dosages led to reduced grinding efficiency, indicating a threshold beyond which additional GA becomes counterproductive.

At 70 min of grinding, the performance benefit of GAs persisted, with Blaine fineness values ranging from 8% to 25% higher than the control sample. This long-term grinding evaluation confirms that all tested GAs contribute positively to grinding productivity, albeit with varying effectiveness depending on the grinding duration and GA chemistry.

Overall, these findings provide a clearer understanding of the quantitative impact of GA selection on cement fineness. Compared to traditional cement grinding without additives, the inclusion of optimized GA types and dosages can improve Blaine fineness by up to 25%, depending on the grinding duration and chemical structure of the additive. These results offer a measurable and comparative basis for selecting suitable GAs to enhance grinding efficiency in industrial applications.

3.2. Evaluation of Modeling Results

Using the results obtained from the experimental data, dominant parameters, determination of the coefficients of the parameters, and modeling studies were carried out using machine and ensemble learning methods of Blaine value, using real and augmented data. Feature importance values for the XGBoost method were used to determine the dominant parameters for grinding efficiency values. The parameters used to determine the dominant parameters with the XGBoost method are presented in Table 3, and the results obtained as a result of the evaluations are presented in Table 4.

Table 3. Parameters used in experiments for the RF method.

Experiment Data	Parameters			
	Estimator	Depth	Learning Rate	Tree Method
Real	47	4	0.3	hist
Augmented	47	4	0.3	exact

Table 4. Parameters affecting grinding efficiency values (feature importance values).

Model	GA Dosage (%)	Molecular Weight	Density	pH	Number of Hydrogen Bond Acceptors	Number of Functional Groups	Grinding Time (s)
Real Data	0.000749	0.005852	0.004484	0.002404	0.000531	0.000001	0.98598
Sorting	5	2	3	4	6	7	1
Augmented Data	0.001604	0.004744	0.003322	0.002009	0.001754	0.000001	0.986568
Sorting	6	2	3	4	5	7	1

The coefficients calculated by the LR method for the grinding efficiency experiment are presented in Table 5. In the table, weights are calculated for seven features (inputs) according to the model created with the augmented data of the grinding efficiency experi-

ment. By using the weights of the features, predictions can be made for data not available in the dataset or available to another researcher. In the grinding efficiency experiment, the weight coefficient of 659.54 or -204.7 for the GA feature was multiplied by the value found in the dataset. After applying this process to the other features, the intercept value in the last column of the table was added to the result. The result obtained was the initial estimate of the data for that experiment.

Table 5. Equationization of the calculated coefficients.

Experiment	Coefficients							
	GA	Molecular Weight	Density	pH	HBA	FGA	GT	Intercept
Blaine(Real Data)	-204.7	-0.53	-1764.46	-36.75	83.67	83.67	40.37	3049.43
Blaine(Augmented Data)	659.54	-0.1	-1192.59	-25.63	24.29	24.29	41.23	2561.88

Results for Real Data

The best values obtained from the models on real data for the grinding efficiency experiment are given in Table 6.

Table 6. Best values obtained from models on real data for the grinding efficiency experiment.

Blaine Model	MAE		RMSE		LogCosh	
	Train	Test	Train	Test	Train	Test
ANN	71.7741	90.4075	91.0063	117.8307	89.7182	71.0810
TabNet	40.6502	98.1536	56.0825	119.0298	39.9612	97.4604
RF	29.1276	65.7865	38.5301	79.8695	28.4370	46.7729
XGBoost	7.1568	39.4859	10.9029	49.0680	22.1620	22.6828

For the grinding efficiency experiment, the results obtained on the real and augmented data with the ANN, TabNet, RF, and XGBoost methods are presented in Tables 6 and 7. The training column shows the results obtained by the method in the learning phase and the test column shows the prediction success of the method in the test phase. The value with the best test result is bolded. As emphasized before, three different metrics, RMSE, MAE, and Logcosh, were used to evaluate the results obtained. As the error metric values approach 0, the prediction success of the method increases, and as they move away, the prediction success decreases. On the real data of the grinding efficiency experiment, the best test results were found to be the XGBoost and RF methods for the RMSE, MAE, and Logcosh metrics, respectively. The TabNet method showed the worst performance in the test results. In terms of metrics, Logcosh achieved the best training and test values in all parameters. This was followed by MAE and RMSE metric results. When the results for three different metrics are examined in all modeling studies, the order of success of the metrics is Logcosh, MAE, and RMSE. In addition to the fact that the modeling results are close to the experimental results, the same ranking of the metric values obtained despite the use of different methods can be presented as evidence of the accuracy of the modeling. The graphical representation of the experimental and model results of the train and test phases with real data is shown in Figures 5 and 6.

Table 7. The best values obtained from the models on the augmented data for the grinding efficiency experiment.

Blaine	MAE		RMSE		LogCosh	
Model	Train	Test	Train	Test	Train	Test
ANN	43.3705	39.9442	52.6161	49.9826	42.6816	39.2522
TabNet	53.9081	52.8663	62.6867	61.9604	53.2175	52.1758
RF	38.1095	42.5909	56.5102	62.9739	39.8402	40.0979
XGBoost	11.2431	21.0384	16.6611	33.7379	10.5850	15.4846

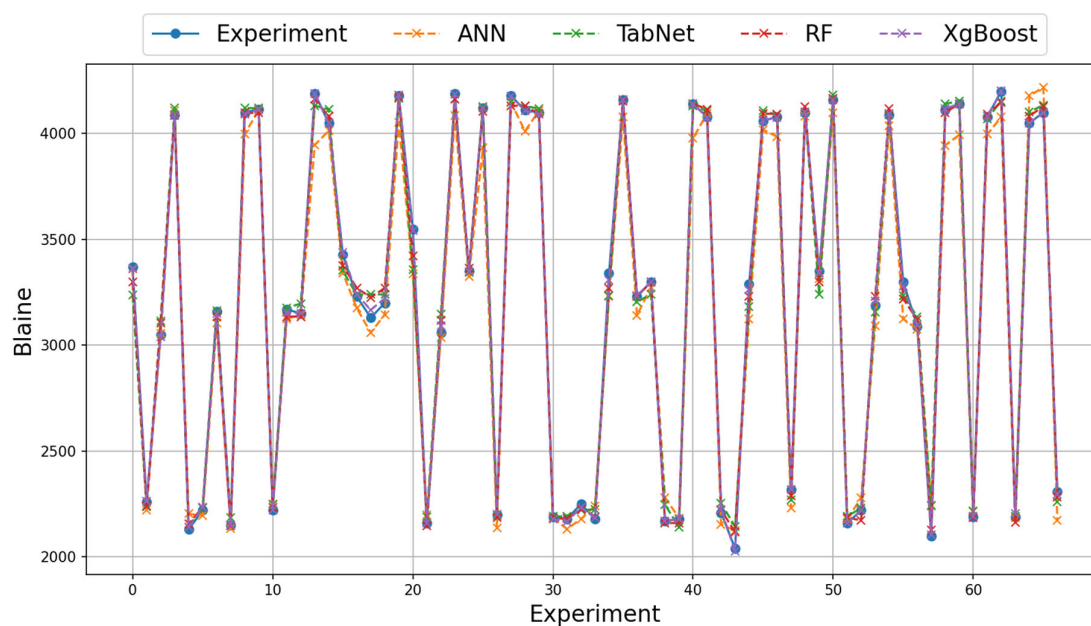


Figure 5. Illustration of experimental and model results on real training data.

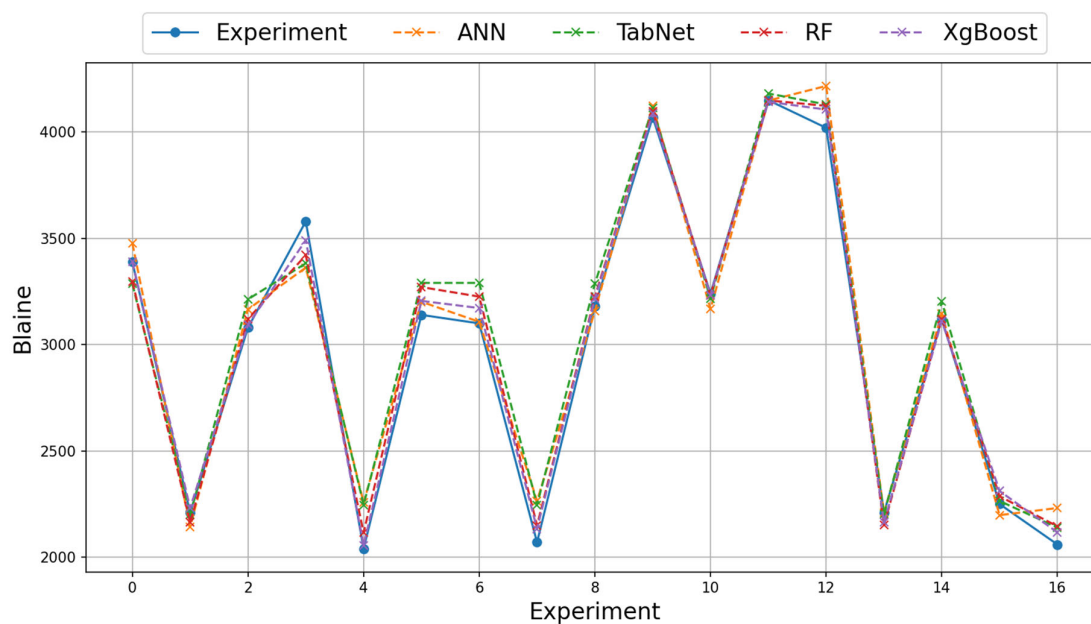


Figure 6. Illustration of experimental and model results on real test data.

On the real data of grinding efficiency experiment, the best test results were found to be the XGBoost, ANN, and RF methods for the RMSE, MAE, Logcosh metrics, respectively. Better results were found in the RMSE metric in the training phase using the ANN

method and in the MAE and Logcosh metrics using the RF method. In the test phase, the ANN method was found to be better than the RF method in all metrics. The TabNet method generally showed the worst performance in the test and train results. In terms of metrics, Logcosh achieved the best training and test values for all parameters. This was followed by the MAE and RMSE metric results. When the results for three different metrics are examined across all modeling studies, the order of success of the metrics is Logcosh, MAE, and RMSE. In addition to the fact that the modeling results are close to the experimental results, the same ranking of the metric values obtained despite the use of different methods can be presented as evidence of the accuracy of the modeling. The graphical representation of the experimental and model results of the train and test phases with real data is shown in Figures 7 and 8.

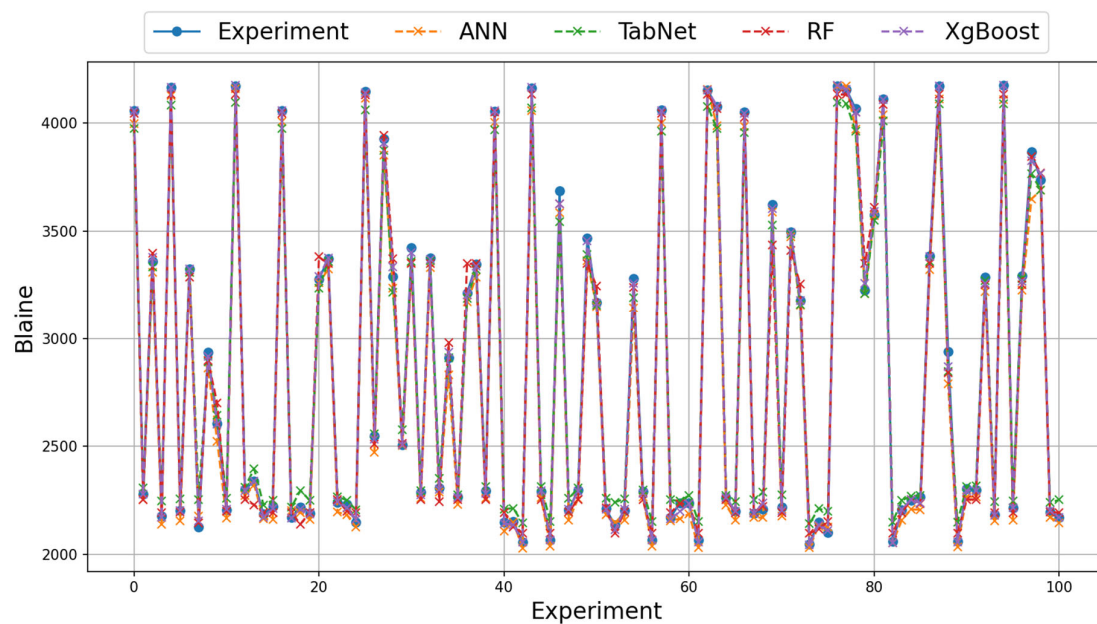


Figure 7. Illustration of experimental and model results on augmented training data (first 100 samples).

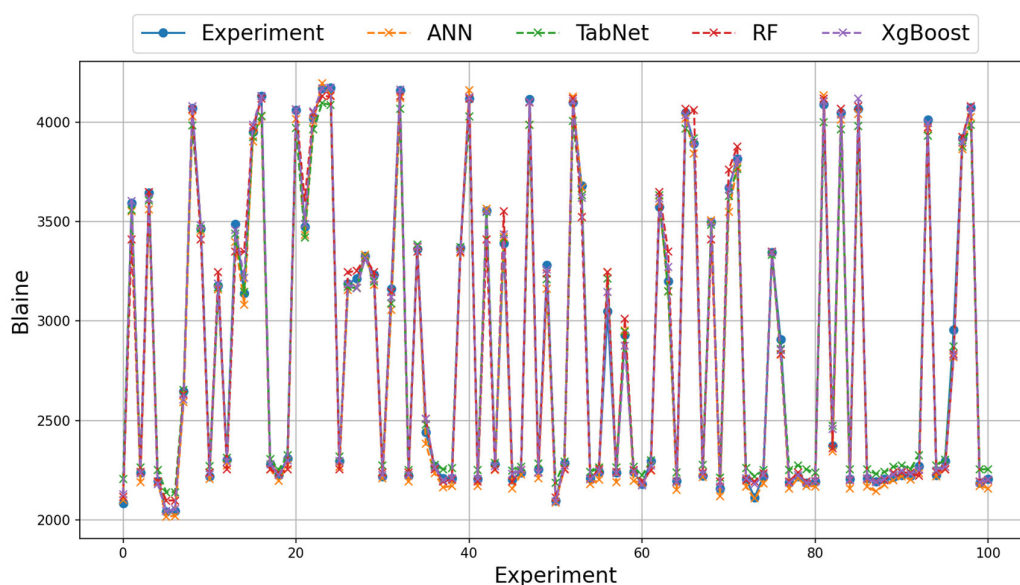


Figure 8. Illustration of experimental and model results on augmented test data (first 100 samples).

4. Conclusions

In this study, Blaine fineness values of cements obtained in the presence of different GA types were modeled using machine learning, transformer-based learning, and ensemble learning methods. The results derived from the obtained data are summarized below, organized under the headings of scientific findings, effective parameters, and application outcomes.

Scientific Findings:

1. The XGBoost method demonstrated superior performance in modeling Blaine fineness values for cements with different grinding aid (GA) types, using both real and augmented datasets. Its efficiency in rapidly generating tree-based models makes it a powerful tool for regression and statistical modeling in cementitious systems.
2. Data augmentation through the Synthetic Minority Over-sampling Technique (SMOTE) significantly improved the predictive accuracy across all modeling methods. Consistency in the ranking of error metrics (Logcosh, MAE, and RMSE) validates the robustness and reliability of the models developed.
3. Ensemble learning methods, such as Random Forest and XGBoost, exhibited strong predictive capability even with limited dataset sizes, whereas Artificial Neural Networks (ANNs) and transformer-based TabNet require larger datasets to achieve comparable success.
4. Parameter importance analysis confirmed grinding time as the most influential factor affecting Blaine fineness, underscoring the physical relevance and accuracy of the modeling approaches.

Effective parameters: milling time, GA dosage, molecular weight, density, pH, number of functional groups, and number of hydrogen bond acceptors.

Applied Results:

1. The integration of advanced machine learning and ensemble methods with data augmentation enables more accurate prediction and optimization of grinding processes in cement production, potentially reducing trial-and-error experimental efforts and accelerating product development cycles.
2. The identification of key parameters influencing grinding efficiency can guide the targeted adjustment of process variables and grinding aid formulations, improving operational efficiency and product quality.

Author Contributions: Methodology, Y.K., V.K., M.E., N.M., M.B. and A.M.; software, Y.K., V.K., M.E., N.M., M.B. and A.M.; investigation, Y.K., V.K., M.E., N.M., M.B. and A.M.; resources, Y.K., V.K., M.E., N.M., M.B. and A.M.; writing—original draft preparation, Y.K., V.K., M.E., N.M., M.B. and A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Bursa Uludağ University BAP, grant number FPDD-2025-2210.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors express their sincere appreciation to the Scientific and Technological Research Council of Turkey (TÜBİTAK) for the support provided under Grant No. 222M425. They also acknowledge the contributions of the Bursa Uludağ University Science and Technology Centre (BAP) under Project No. FPDD-2025-2210. The first author gratefully acknowledges TÜBİTAK for the 2211A scholarship awarded during their doctoral studies. Additionally, the last author extends their heartfelt gratitude to the Turkish Academy of Sciences (TÜBA) for their valuable support.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kobya, V.; Karakuzu, K.; Mardani, A.; Felekoğlu, B.; Ramyar, K. Combined interaction of PCE chains lengths, C3A and water content in cementitious systems. *Constr. Build. Mater.* **2023**, *378*, 131178. [\[CrossRef\]](#)
2. Şahin, H.G.; Biricik, Ö.; Mardani-Aghabaglou, A. Polycarboxylate-based water reducing admixture–clay compatibility; literature review. *J. Polym. Res.* **2022**, *29*, 33. [\[CrossRef\]](#)
3. Lei, L.; Hirata, T.; Plank, J. 40 years of PCE superplasticizers-History, current state-of-the-art and an outlook. *Cem. Concr. Res.* **2022**, *157*, 106826. [\[CrossRef\]](#)
4. Ma, Y.; Sha, S.; Zhou, B.; Lei, F.; Liu, Y.; Xiao, Y.; Shi, C. Adsorption and dispersion capability of polycarboxylate-based superplasticizers: A review. *J. Sustain. Cem.-Based Mater.* **2022**, *11*, 319–344. [\[CrossRef\]](#)
5. Özen, S.; Altun, M.G.; Mardani-Aghabaglou, A.; Ramyar, K. Effect of main and side chain length change of polycarboxylate-ether-based water-reducing admixtures on the fresh state and mechanical properties of cementitious systems. *Struct. Concr.* **2021**, *22*, E607–E618. [\[CrossRef\]](#)
6. Sha, S.; Wang, M.; Shi, C.; Xiao, Y. Influence of the structures of polycarboxylate superplasticizer on its performance in cement-based materials-A review. *Constr. Build. Mater.* **2020**, *233*, 117257. [\[CrossRef\]](#)
7. Yiğit, B.; Salihoğlu, G.; Mardani-Aghabaglou, A.; Salihoğlu, N.K.; Özen, S. Recycling of sewage sludge incineration ashes as construction material. *J. Fac. Eng. Archit. Gazi Univ.* **2020**, *35*, 1647–1664.
8. Aytekin, B.; Mardani-Aghabaglou, A. Sustainable materials: A review of recycled concrete aggregate utilization as pavement material. *Transp. Res. Rec.* **2022**, *2676*, 468–491. [\[CrossRef\]](#)
9. Giergiczny, Z. Fly ash and slag. *Cem. Concr. Res.* **2019**, *124*, 105826. [\[CrossRef\]](#)
10. Mathapati, M.; Amate, K.; Prasad, C.D.; Jayavardhana, M.L.; Raju, T.H. A review on fly ash utilization. *Mater. Today Proc.* **2022**, *50*, 1535–1540. [\[CrossRef\]](#)
11. Becerra-Duitama, J.A.; Rojas-Avellaneda, D. Pozzolans: A review. *Eng. Appl. Sci. Res.* **2022**, *49*, 495–504.
12. Chipakwe, V.; Semsari, P.; Karlkvist, T.; Rosenkranz, J.; Chelgani, S.C. A critical review on the mechanisms of chemical additives used in grinding and their effects on the downstream processes. *J. Mater. Res. Technol.* **2020**, *9*, 8148–8162. [\[CrossRef\]](#)
13. Hao, S.H.; Liu, B.H.; Yan, X.Y. Review on research of cement grinding aids and certain problems. *Key Eng. Mater.* **2017**, *753*, 295–299. [\[CrossRef\]](#)
14. Nthiga Njiru, E.; Wachira Muthengia, J.; Mulwa Munyao, O.; Karanja Mutitu, D.; Munyao Musyoki, D. Review of the Effect of Grinding Aids and Admixtures on the Performance of Cements. *Adv. Civ. Eng.* **2023**, *2023*, 6697842. [\[CrossRef\]](#)
15. Assaad, J.J.; Asseily, S.E.; Harb, J. Use of cement grinding aids to optimise clinker factor. *Adv. Cem. Res.* **2010**, *22*, 29–36. [\[CrossRef\]](#)
16. Prziwara, P.; Breitung-Faes, S.; Kwade, A. Impact of grinding aids on dry grinding performance, bulk properties and surface energy. *Adv. Powder Technol.* **2018**, *29*, 416–425. [\[CrossRef\]](#)
17. Assaad, J.J.; Issa, C.A. Effect of clinker grinding aids on flow of cement-based materials. *Cem. Concr. Res.* **2014**, *63*, 1–11. [\[CrossRef\]](#)
18. Kaya, Y.; Kobya, V.; Mardani, A.; Assaad, J.J. Effect of Modified Triethanolamine on Grinding Efficiency and Performance of Cementitious Materials. *Talanta Open* **2024**, *9*, 100293. [\[CrossRef\]](#)
19. Durgun, M.Y.; Özen, S.; Karakuzu, K.; Kobya, V.; Bayqra, S.H.; Mardani-Aghabaglou, A. Effect of high temperature on polypropylene fiber-reinforced mortars containing colemanite wastes. *Constr. Build. Mater.* **2022**, *316*, 125827. [\[CrossRef\]](#)
20. Jolicœur, J.; Morasse, S.; Sharman, J.; Tagnit-Hamou, A.; Slim, F.; Page, M. Polyol-type compounds as clinker grinding aids: Influence of powder fluidity and on cement hydration. In Proceedings of the 12th International Congress on the Chemistry of Cement, Montréal, QC, Canada, 8–13 July 2007; pp. 1–6.
21. Zhang, T.; Gao, J.; Hu, J. Preparation of polymer-based cement grinding aid and their performance on grindability. *Constr. Build. Mater.* **2015**, *75*, 163–168. [\[CrossRef\]](#)
22. Sun, Z.; Liu, H.; Ji, Y.; Pang, M. Influence of glycerin grinding aid on the compatibility between cement and polycarboxylate superplasticizer and its mechanism. *Constr. Build. Mater.* **2020**, *233*, 117104. [\[CrossRef\]](#)
23. Kobya, V.; Kaya, Y.; Mardani-Aghabaglou, A. Effect of amine and glycol-based grinding aids utilization rate on grinding efficiency and rheological properties of cementitious systems. *J. Build. Eng.* **2022**, *47*, 103917. [\[CrossRef\]](#)
24. Zhao, J.; Wang, D.; Wang, X.; Liao, S. Characteristics and mechanism of modified triethanolamine as cement grinding aids. *J. Wuhan Univ. Technol.-Mater. Sci. Ed.* **2015**, *30*, 134–141. [\[CrossRef\]](#)
25. Kaya, Y.; Kobya, V.; Mardani, A. Evaluation of fresh state, rheological properties, and compressive strength performance of cementitious system with grinding aids. *J. Appl. Polym. Sci.* **2024**, *141*, e55212. [\[CrossRef\]](#)
26. Sutton, R.S.; Barto, A. Reinforcement learning: An introduction. *IEEE Trans. Neural Netw.* **1998**, *9*, 1054. [\[CrossRef\]](#)
27. Mardani-Aghabaglou, A.; Özen, S.; Altun, M.G. Durability performance and dimensional stability of polypropylene fiber reinforced concrete. *J. Green Build.* **2018**, *13*, 20–41. [\[CrossRef\]](#)

28. Ai, Y.; Ye, C.; Liu, J.; Zhou, M. Study on the evolution processes of keyhole and melt pool in different laser welding methods for dissimilar materials based on a novel numerical model. *Int. Commun. Heat Mass Transf.* **2025**, *163*, 108629. [\[CrossRef\]](#)
29. Khan, I.U.; Ouaisa, M.; Ouaisa, M.; Fayaz, M.; Ullah, R. (Eds.) *Artificial Intelligence for Intelligent Systems: Fundamentals, Challenges, and Applications*; CRC Press: Boca Raton, FL, USA, 2024.
30. Zhao, Z.; Fan, W.; Li, J.; Liu, Y.; Mei, X.; Wang, Y.; Wen, Z.; Wang, F.; Zhao, X.; Tang, J.; et al. Recommender systems in the era of large language models (llms). *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 6889–6907. [\[CrossRef\]](#)
31. Bouguettaya, A.; Zarzour, H.; Kechida, A.; Taberkit, A.M. Machine learning and deep learning as new tools for business analytics. In *Handbook of Research on Foundations and Applications of Intelligent Business Analytics*; IGI Global Scientific Publishing: Hershey, PA, USA, 2022; pp. 166–188.
32. Hosni, M.; Abnane, I.; Idri, A.; de Gea, J.M.C.; Alemán, J.L.F. Reviewing ensemble classification methods in breast cancer. *Comput. Methods Programs Biomed.* **2019**, *177*, 89–112. [\[CrossRef\]](#)
33. Yüksel, C.; Mardani-Aghabaglou, A.; Beglarigale, A.; Yazıcı, H.; Ramyar, K.; Andiç-Çakır, Ö. Influence of water/powder ratio and powder type on alkali-silica reactivity and transport properties of self-consolidating concrete. *Mater. Struct.* **2016**, *49*, 289–299. [\[CrossRef\]](#)
34. Fatahi, R.; Pournazari, A.; Ghorbani, H.; Shah, M.P.; tazik, S.H.; Abdollahi, H. Modeling and optimization of pea- combined grinding aids effect in grinding circuit of cement ball mill. *Res. Sq.* **2023**. [\[CrossRef\]](#)
35. Chen, L.Q.; Guo, J.L.; Yang, X.; Chi, J.; Zhao, X. Grinding roughness prediction model based on evolutionary artificial neural network. *Comput. Integr. Manuf. Syst.* **2013**, *19*.
36. Karkalos, N.E.; Markopoulos, A.P.; Dossis, M.F. Application of statistical and soft computing techniques for the prediction of grinding performance. *J. Robot Mech. Eng. Resr.* **2015**, *1*, 6–16.
37. Pawuś, D.; Paszkiel, S. Identification and Expert Approach to Controlling the Cement Grinding Process Using Artificial Neural Networks and Other Non-Linear Models. *IEEE Access* **2024**, *12*, 26364–26383. [\[CrossRef\]](#)
38. Liu, Z.; Wang, X.; Zhang, Q.; Huang, C. Empirical mode decomposition based hybrid ensemble model for electrical energy consumption forecasting of the cement grinding process. *Measurement* **2019**, *138*, 314–324. [\[CrossRef\]](#)
39. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [\[CrossRef\]](#)
40. EN 197-1; Cement: Composition, Specifications and Conformity Criteria, Part 1: Common Cements. European Committee for Standardization: Bruxelles, Belgium, 2000; EN/TC51/WG 6 Rev.
41. Dake, D.K.; Gyimah, E. Using sentiment analysis to evaluate qualitative students' responses. *Educ. Inf. Technol.* **2022**, *28*, 4629–4647. [\[CrossRef\]](#)
42. Ramezan, C.A.; Warner, T.A.; Maxwell, A.E. Evaluation of sampling and cross-validation tuning strategies for regional-scale machine learning classification. *Remote Sens.* **2019**, *11*, 185. [\[CrossRef\]](#)
43. Krstajic, D.; Buturović, L.; Leahy, D.E.; Thomas, S. Cross-validation pitfalls when selecting and assessing regression and classification models. *J. Cheminformatics* **2014**, *6*, 10. [\[CrossRef\]](#)
44. Poudyal, R.; Paneru, B.; Paneru, B.; Giri, T.; Paneru, B.; Reynolds, T.; Poudyal, K.N.; Dangi, M.B. Exploring Cement Production's Role in GDP Using Explainable AI and Sustainability Analysis in Nepal. *Case Stud. Chem. Environ. Eng.* **2025**, *11*, 101128. [\[CrossRef\]](#)
45. Zhang, A.; Yu, H.; Huan, Z.; Yang, X.; Zheng, S.; Gao, S. SMOTE-RkNN: A hybrid re-sampling method based on SMOTE and reverse k-nearest neighbors. *Inf. Sci.* **2022**, *595*, 70–88. [\[CrossRef\]](#)
46. Barratt, S.; Sharma, R. Optimizing for generalization in machine learning with cross-validation gradients. *arXiv* **2018**, arXiv:1805.07072. [\[CrossRef\]](#)
47. Granholm, V.; Noble, W.S.; Käll, L. A cross-validation scheme for machine learning algorithms in shotgun proteomics. *BMC Bioinformatics* **2012**, *13* (Suppl. 16), S3. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Mardani-Aghabaglou, A.; Öztürk, H.T.; Kankal, M.; Ramyar, K. Assessment and prediction of cement paste flow behavior; Marsh-funnel flow time and mini-slump values. *Constr. Build. Mater.* **2021**, *301*, 124072. [\[CrossRef\]](#)
49. Mardani-Aghabaglou, A.; Kankal, M.; Nacar, S.; Felekoğlu, B.; Ramyar, K. Assessment of cement characteristics affecting rheological properties of cement pastes. *Neural Comput. Appl.* **2021**, *33*, 12805–12826. [\[CrossRef\]](#)
50. Montgomery, D.C.; Peck, E.A.; Vining, G.G. *Introduction to Linear Regression Analysis*, 5th ed.; Wiley: Hoboken, NJ, USA, 2012.
51. Assaad, J.J.; Asseily, S.E.; Harb, J. Effect of specific energy consumption on fineness of portland cement incorporating amine or glycol-based grinding aids. *Mater. Struct.* **2009**, *42*, 1077–1087. [\[CrossRef\]](#)
52. Mishra, R.K.; Weibel, M.; Müller, T.; Heinz, H.; Flatt, R.J. Energy-effective grinding of inorganic solids using organic additives. *Chim. Int. J. Chem.* **2017**, *71*, 451–460. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.