



Comparación de los modelos LightGBM y XGBoost para la predicción a la resistencia a la compresión del concreto

Comparison of LightGBM and XGBoost models for predicting concrete compressive strength

Comparação dos modelos LightGBM e XGBoost para previsão da resistência à compressão do concreto

ARTÍCULO ORIGINAL

Jaime Yelsin Rosales Malpartida

jrosalesm@uni.pe

<https://orcid.org/0000-0003-4574-5172>

Facultad de Ingeniería Civil

Universidad Nacional de Ingeniería, Lima, Perú

Victor Adrian Rosales-Cutipa

victor.rosales.c@uni.pe

<https://orcid.org/0009-0004-5651-3726>

Facultad de Ingeniería Civil

Universidad Nacional de Ingeniería, Lima, Perú

Recibido 02 de Mayo 2025 | Arbitrado y aceptado 02 de Junio 2025 | Publicado el 16 de Julio 2025

RESUMEN

El uso de modelos de Machine Learning (ML), particularmente aquellos basados en gradient-boosting, está transformando la predicción de propiedades materiales con relaciones no lineales complejas. Este artículo presenta un análisis exhaustivo del desempeño de dichos modelos en la predicción de la resistencia a la compresión del concreto en función de las características del diseño de mezcla utilizando 525 datos experimentales obtenidos del registro histórico del Laboratorio N°1 de Ensayo de Materiales "Ing. Manuel Gonzales de la Cotera" de la Facultad de Ingeniería Civil de la Universidad Nacional de Ingeniería (LEM), correspondientes al período 2015-2022. La base de datos se utilizó para entrenar y probar dos modelos de ML optimizados los cuales fueron Light GBM y XGBoost. Los modelos se entrenaron utilizando la optimización de Grid Search para ajustar los hiperparámetros de configuración y obtener el mejor rendimiento de para ambos modelos. Los resultados muestran que el modelo XGBoost tiene mayor rendimiento tanto en la fase de entrenamiento como en la de prueba en comparación al modelo Light GBM en cuanto a los parámetros estadísticos como RMSE, MAE, MAPE y R^2 .

Palabras Clave: Machine Learning, Gradient-boosting, diseño de mezcla, modelos y Grid Search.

ABSTRACT

The use of Machine Learning (ML) models, particularly those based on gradient-boosting, is transforming the prediction of material properties with complex non-linear relationships. This paper presents a comprehensive analysis of the performance of such models in predicting the compressive strength of concrete based on mix design characteristics using 525 experimental data obtained from the historical record of the Materials Testing Laboratory No. 1 "Ing. Manuel Gonzales de la Cotera" of the Faculty of Civil Engineering of the National University of Engineering (LEM), corresponding to the period 2015-2022. The database was used to train and test two optimized ML models which were Light GBM and XGBoost. The models were trained using Grid Search optimization to tune the configuration hyperparameters and obtain the best performance for both models. The results show that the XGBoost model has a higher performance in both the training and testing phases compared to the Light GBM model in terms of statistical parameters such as RMSE, MAE, MAPE and R^2 .

Keywords: Machine Learning, Gradient-boosting, mixture design, models and Grid Search.

RESUMO

O uso de modelos de aprendizado de máquina (ML), particularmente aqueles baseados em aumento de gradiente, está transformando a previsão de propriedades de materiais com relacionamentos não lineares complexos. Este artigo apresenta uma análise abrangente do desempenho de tais modelos na previsão da resistência à compressão do concreto com base nas características do projeto da mistura usando 525 dados experimentais obtidos do registro histórico do Laboratório de Ensaio de Materiais n° 1 "Ing. Manuel Gonzales de la Cotera" da Faculdade de Engenharia Civil da Universidade Nacional de Engenharia (LEM), correspondente ao período de 2015-2022. O banco de dados foi usado para treinar e testar dois modelos de ML otimizados, que eram Light GBM e XGBoost. Os modelos foram treinados usando a otimização do Grid Search para ajustar os hiperparâmetros de configuração e obter o melhor desempenho para ambos os modelos. Os resultados mostram que o modelo XGBoost tem um desempenho superior tanto na fase de treinamento quanto na de teste em comparação ao modelo Light GBM em termos de parâmetros estatísticos como RMSE, MAE, MAPE e R^2 .

Palavras-chave: Aprendizado de máquina, aumento de gradiente, design de mistura, modelos e busca em grade.

1. INTRODUCCIÓN

Uno de los materiales más usados en la industria de la construcción es el concreto [1]. Este material resulta de la mezcla de diferentes componentes, como el cemento, agua, agregados, aditivos y adiciones [2]. La construcción se beneficia ampliamente de la versatilidad del concreto, tanto en su estado convencional como reforzado, gracias a sus propiedades y a su capacidad de combinarse con materiales de refuerzo [3].

Los componentes del concreto y sus proporciones guardan una estrecha relación con la resistencia a la compresión que este puede alcanzar [3], [4], [5], [6]. La tecnología del concreto tiende a emplear esta propiedad como principal indicador de la calidad del material, así como indicador cuantitativo para otras propiedades del concreto endurecido [2], [3].

La resistencia a la compresión es un parámetro esencial para garantizar la seguridad, durabilidad y desempeño de las estructuras [3]. El estudio de esta propiedad es de gran importancia, y el proceso por el cual se realiza es tedioso y tardado [3], [7], [8]. Además, diversos factores del proceso, las particularidades de los componentes y la aleatoriedad de las pruebas afectan los resultados [7], [9], [10], requiriendo así una presencia significativa de empirismo y experiencia en los diseños más complejos [11], [12].

El aprendizaje automático se ha convertido en una herramienta importante para la predicción de propiedades que mantienen relaciones no lineales complejas con sus parámetros [13], [14], [15]. Esta capacidad ha sido empleada previamente en el campo de la ciencia de los materiales [16], [17], [18]. Diversas investigaciones han estudiado la predicción de propiedades mecánicas del concreto utilizando modelos como redes neuronales artificiales (ANN), árboles de decisión (DT) y máquinas de soporte vectorial (SVM) [10], [19], [20], [21].

Dos novedosos algoritmos han sido foco de atención debido a su fiabilidad y eficacia corroborada en investigaciones [1], [22], [23], [24], [25]. XGBoost y LightGBM destacan por ser algoritmos rápidos y eficientes caracterizados por su metodología de gradient-boosting [13], [23], [26], [27]. Su correcta implementación permite que sus resultados de predicción sobresalgan sobre los de otros algoritmos

gracias a su excepcional rendimiento [1]. Este destacado desempeño, combinado con su alta explicabilidad, ha propiciado una amplia adopción de estos algoritmos en la ingeniería civil [23].

Este artículo presenta una comparación de los algoritmos previamente mencionados y sus desempeños al emplearse en la predicción de la resistencia a la compresión de muestras cilíndricas de concreto a partir de sus proporciones de diseño y características de sus componentes.

2. ANTECEDENTES

En la revisión de literatura se encuentran diversos estudios orientados a la evaluación de algoritmos de predicción para la obtención de propiedades del concreto. Los modelos ANN, DT y SVM han sido ampliamente empleados en investigaciones previas en conjunto con algoritmos particulares [21].

Respecto a la resistencia a la compresión, el estudio de esta propiedad ha sido realizado desde diversos enfoques como los modelos matemáticos, estadísticos, analíticos, numéricos y métodos computacionales [28]. El ensamblaje de algoritmos también es una técnica empleada por autores a fin de obtener mejores resultados en torno a la predicción de propiedades [16], [19], [29].

Un aspecto importante para considerar en los modelos de predicción supervisada son los parámetros de entrada y salida [30]. El contenido de agua, cemento, agregado grueso, agregado fino, aditivo y el tiempo de curado, son parámetros recurrentemente encontrados en la literatura [19], [31], [32], [33], [34]. Ocasionalmente, se identifican estudios que consideran propiedades específicas de los agregados, como los módulos de finura, pesos específicos o porcentajes de absorción [19], [35]; así como también, algunos consideran porciones específicas del contenido de agregado fino, usualmente diferenciadas por la malla que retiene a estas porciones [10].

Estudios equivalentes que emplean los algoritmos gradient-boosting arrojan resultados que destacan las capacidades particulares de estos modelos [13]. Comparaciones indican un mejor desempeño del algoritmo LightGBM al aplicarlo para la predicción de la resistencia a la penetración rápida de cloruros en concreto de alta resistencia, destacando en comparación a XGBoost [27]. Por otra parte, autores que

emplean estos métodos en la resolución de problemas de otra índole obtienen resultados contrarios, como el caso de aplicarlo a la predicción de la resistencia a la compresión residual del concreto expuesto a elevadas temperaturas [36]. Ejemplificando que los métodos pueden presentar un mejor desempeño según sean las condiciones en las que se implementa, involucrando al contexto, la naturaleza del problema y los parámetros de los cuales se entrena y prueba.

Este estudio tiene como objetivo emplear y evaluar el desempeño de los algoritmos de gradient-boosting en la predicción de la resistencia a la compresión del concreto, utilizando datos históricos recopilados en un laboratorio de ensayo de materiales. La investigación busca contribuir al conocimiento sobre el uso de estos algoritmos innovadores en una disciplina aún no completamente explorada, aprovechando los últimos avances en algoritmos de predicción.

3. METODOLOGÍA

En esta sección se describen los aspectos clave de la metodología empleada en la investigación, incluyendo la base de datos experimental utilizada, su modo de obtención, los algoritmos de predicción empleados y los indicadores para realizar la comparación entre sus desempeños. Se detalla cómo se estructuran y analizan los datos para alcanzar los objetivos del estudio.

3.1 BASE DE DATOS EXPERIMENTAL

Se recopilaron un total de 525 sets de datos experimentales provenientes del registro histórico de ensayos de diseño de mezcla del Laboratorio N°1 de Ensayo de Materiales “Ing. Manuel Gonzales de la Cotera” de la facultad de Ingeniería Civil de la Universidad Nacional de Ingeniería (LEM en adelante). Aquellos datos, correspondientes a ensayos realizados durante el periodo 2015-2022.

Respecto a la selección de los ensayos, se priorizaron aquellos que incluían solo un aditivo o ninguno. Además, se excluyeron los ensayos pertenecientes a proyectos de investigación (tesis), dando prioridad a los diseños empleados en construcciones civiles, independientemente del tipo de proyecto. De esta manera, se contemplaron concretos utilizados para pavimentos, edificaciones, cimentaciones y otros usos no estructurales.

El ensayo de diseño de mezcla realizado en el LEM contempla la realización de ensayos auxiliares para obtener las propiedades de los agregados que compondrán el concreto, así como también, contempla el ensayo de compresión de probetas cilíndricas, destinado a corroborar la calidad del diseño realizado. Por tanto, la cantidad de parámetros posibles de obtener es muy alta. Los parámetros han sido clasificados en secciones de acuerdo con su naturaleza, siendo estas las dedicadas a datos administrativos, datos de cemento, aditivo, agregado grueso, agregado fino, mezcla y probeta endurecida.

Cabe destacar que, debido a la naturaleza del ensayo de diseño de mezcla, un mismo diseño requiere de dos a cuatro probetas para verificar la resistencia a la compresión alcanzada. Cada una de esas probetas ensayadas da pie a la creación de un set de datos experimentales.

La preparación de los parámetros de entrada contempló la creación de seis adicionales: la relación agua-cemento, el agua adicional, el agua real, la relación agua-cemento real, la relación agregado fino-cemento y la relación agregado grueso-cemento. Considerando estas inclusiones, la presente investigación emplea un total de 50 variables, incluyendo a la resistencia a la compresión.

Aunque algunos autores sugieren que un mayor número de variables de entrada puede mejorar la precisión de las predicciones [10], un exceso de parámetros también puede provocar problemas de sobreajuste en los modelos [16], [37]. No obstante, se espera que los algoritmos seleccionados puedan manejar esta dificultad gracias a sus características específicas. Así, se propone estimar un único valor: la resistencia a la compresión de probetas de concreto.

3.2 MODELOS GRADIENT-BOOSTING

El gradient-boosting es una técnica de aprendizaje automático que construye un modelo predictivo fuerte a partir de múltiples modelos débiles, generalmente árboles de decisión. Funciona iterativamente, optimizando el modelo al corregir los errores de los modelos anteriores mediante la minimización del gradiente de una función de pérdida [38]. Este enfoque permite manejar datos complejos y mejorar la precisión predictiva, haciéndolo especialmente útil en tareas de regresión y clasificación [39]. En este

estudio, se emplean dos novedosos y robustos algoritmos basados en modelos gradient-boosting: LightGBM y XGBoost.

El algoritmo LightGBM, basado en árboles de decisión, incorpora dos estrategias novedosas: el muestreo unilateral basado en gradientes (GOSS) y la agrupación exclusiva de características (EFB) [40]. Estas estrategias mejoran la velocidad de entrenamiento, la precisión y la eficiencia de LightGBM en comparación con los árboles de decisión de gradient-boosting tradicional (GBDT) [38]. La estrategia GOSS permite eliminar una porción significativa de muestras con gradientes pequeños, mejorando así la eficiencia en el cálculo de la ganancia de información [39]. EFB reduce el número de características agrupando las que son mutuamente excluyentes, optimizando el rendimiento del modelo [40].

Por otro lado, el algoritmo XGBoost destaca por su capacidad para manejar datos faltantes y su implementación de técnicas de regularización que previenen el sobreajuste, mejorando la generalización del modelo [41]. Además, XGBoost emplea una optimización paralela que acelera significativamente el proceso de entrenamiento, lo que resulta en una mayor eficiencia computacional [38]. Estas características hacen que XGBoost sea especialmente adecuado para tareas de predicción en conjuntos de datos complejos, proporcionando resultados precisos y confiables [39].

3.3 DESARROLLO DE LOS MODELOS

Aunque los modelos de aprendizaje automático operan según principios y teorías propias, su rendimiento puede ser ajustado a través de la configuración de hiperparámetros. Sin embargo, encontrar los hiperparámetros óptimos para un problema específico no es una tarea sencilla si se sigue un enfoque de prueba y error [16]. Por ello, en esta investigación se aplicará un método específico diseñado para la optimización de estos valores, con el fin de maximizar la eficiencia del modelo.

La optimización de los hiperparámetros en ambos modelos gradient-boosting se lleva a cabo utilizando el método Grid Search. Durante este proceso, se aparta un subconjunto de datos que se mantiene oculto durante el entrenamiento y se utiliza posteriormente como grupo de validación [27]. Esto permite afinar la precisión de los modelos de aprendizaje automático y mitigar el riesgo de sobreajuste.

Los hiperparámetros a optimizar por cada modelo son presentados en la Tabla I para el modelo LightGBM y en la Tabla II para el modelo XGBoost.

TABLA I

Hiperparámetros de LightGBM para predecir la resistencia a la compresión del concreto

Parámetro	Descripción
learning_rate	Para reducir el paso de gradiente
max_depth	Profundidad del árbol
n_estimators	Construir el máximo número de árboles posible
colsample_bytree	Proporción de submuestra de columnas al construir cada árbol
num_leaves	Número máximo de hojas

TABLA II

Hiperparámetros de XGBoost para predecir la resistencia a la compresión del concreto

Parámetro	Descripción
learning_rate	Para reducir el paso de gradiente
max_depth	Profundidad del árbol
n_estimators	Construir el máximo número de árboles posible
min_child_weight	Número mínimo de instancias para cada nodo
colsample_bytree	Proporción de submuestra de columnas al construir cada árbol

3.4 EVALUACIÓN DE MODELOS

Los modelos desarrollados se evalúan mediante índices, o parámetros, estadísticos en concordancia con la literatura previa [1], [16], [27]. Se emplea el coeficiente de determinación (R^2), el error medio absoluto (MAE), el error de porcentaje medio absoluto (MAPE) y la raíz del error cuadrático medio (RMSE). Las formulaciones matemáticas son descritas en la Tabla III, donde y_i y y^i son, respectivamente, los resultados reales y los predichos por el modelo. Además, y_m representa el valor promedio de los resultados reales. El número de muestras se denota por N .

TABLA III

Representación matemática de los parámetros estadísticos para evaluación de desempeño de algoritmos

Parámetro estadístico	Fórmula
Coeficiente de determinación (R^2)	$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - y^i)^2}{\sum_{i=1}^N (y_i - y_m)^2}$
Error medio absoluto (MAE)	$MAE = \frac{\sum_{i=1}^N y_i - y^i }{N}$
Error de porcentaje medio absoluto (MAPE)	$MAPE = \frac{\sum_{i=1}^N \left \frac{y_i - y^i}{y_i} \right }{N} \times 100$
Raíz del error cuadrático medio (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^N (y^i - y_i)^2}{N}}$

Nota: El valor ideal para R^2 es 1, para el resto de los indicadores, 0.

4. ANÁLISIS DE RESULTADOS

En esta sección se describen los resultados obtenidos del procesamiento analítico de los modelos. Los parámetros estadísticos antes presentados serán tratados como métricas de rendimiento.

4.1 EVALUACIÓN ESTADÍSTICA

Los hiperparámetros fueron ajustados y optimizados para predecir la resistencia a la compresión del concreto en ambos modelos. Los resultados son presentados en la Tabla IV para el modelo LightGBM y en la Tabla V para el modelo XGBoost.

TABLA IV

Hiperparámetros optimizados de LightGBM para predecir la resistencia a la compresión del concreto

Parámetro	Valor	Rango
learning_rate	0.01	0.01 – 0.1
max_depth	3	1 – 8
n_estimators	800	100 – 800
colsample_bytree	0.1	0.1 – 0.70
num_leaves	8	1 - 8

TABLA V

Hiperparámetros optimizados de XGBoost para predecir la resistencia a la compresión del concreto

Parámetro	Valor	Rango
learning_rate	0.01	0.01 – 0.1
max_depth	3	1 – 8
n_estimators	800	100 – 800
min_child_weight	3	1 – 5
colsample_bytree	0.1	0.1 – 0.7

La Tabla VI muestra las métricas de rendimiento (RMSE, MAE y MAPE) para la resistencia a la compresión del concreto utilizando los modelos LightGBM y XGBoost, tanto en el conjunto de entrenamiento como en el de prueba. Los resultados indican que el modelo XGBoost supera consistentemente al modelo LightGBM en todas las métricas evaluadas. En ambos conjuntos de datos, el XGBoost presenta valores menores de RMSE, MAE y MAPE en comparación con LightGBM. Esto sugiere que el modelo XGBoost ofrece un mejor rendimiento general, con menores errores en la predicción de la resistencia a la compresión del concreto.

TABLA VI

Comparación entre el rendimiento del modelo LightGBM y XGBoost para predecir la resistencia a la compresión del concreto.

Fase	Métricas de rendimiento	LightGBM	XGBoost
Entrenamiento	RMSE	12.21040	12.20144
	MAE	8.70207	8.65082
	MAPE	3.86162	3.85130
	R^2	0.9721	0.9722
Prueba	RMSE	16.70301	16.63162
	MAE	12.99431	12.80699
	MAPE	6.08226	6.05480
	R^2	0.9506	0.9521

4.2 PENDIENTES DE REGRESIÓN

El rendimiento de los modelos desarrollados se evaluó mediante el análisis de la pendiente de la regresión lineal, trazando los resultados experimentales en el eje x y los resultados predichos en el eje y (Fig. 1). Este enfoque para evaluar modelos de inteligencia artificial ha sido utilizado previamente por varios investigadores [42], [43], [44]. Según investigaciones previa, una pendiente superior a 0.80 indica una buena concordancia entre los resultados experimentales y los predichos [42].

Los resultados de los modelos predichos fueron comparados con los resultados experimentales, como se muestra en la Fig. 1, para la predicción de la resistencia a la compresión del concreto. Los dos modelos demostraron una fuerte correlación con los resultados experimentales, con pendientes de 0.9697 y 0.9632 en la línea de regresión para el conjunto de entrenamiento de LightGBM y XGBoost, respectivamente. Asimismo, los conjuntos de datos de prueba también mostraron una fuerte correlación en las pendientes de la línea de regresión, en comparación con la pendiente de ajuste ideal (pendiente = 1), como se observa en la Figura X4. En los datos de prueba, se obtuvieron pendientes positivas de 0.9316 y 0.9200 para LightGBM y XGBoost, respectivamente. Dado que las pendientes de regresión superan el valor de 0.8, se concluye que ambos modelos presentan una fuerte correlación entre los resultados experimentales y los predichos.

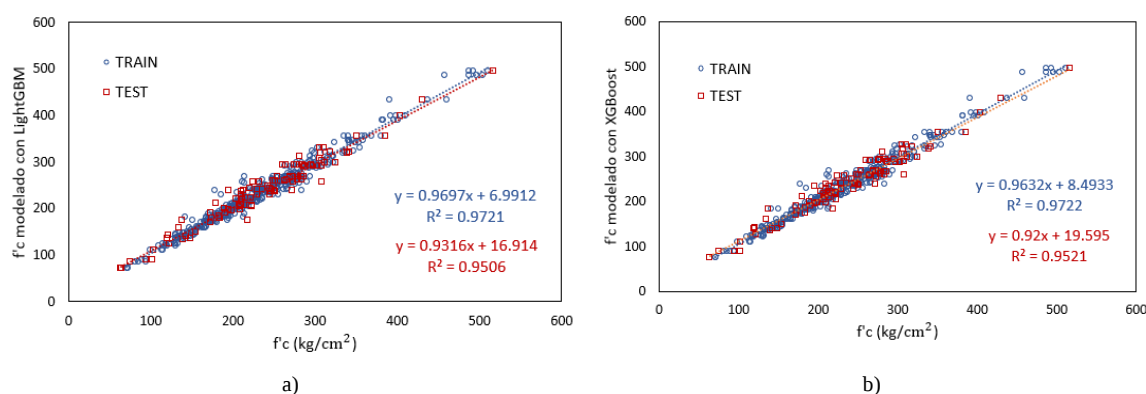


Fig. 1. Comparación de pendientes de regresión de los modelos desarrollados para la resistencia a la compresión del concreto. (a) Pendiente de regresión para el modelo LightGBM, (b) Pendiente de regresión para el modelo XGBoost.

CONCLUSIONES

- El rendimiento óptimo para la predicción de la resistencia a la compresión del concreto se alcanzó con el modelo XGBoost utilizando una tasa de aprendizaje

de 0.01, una profundidad máxima de 8, 800 árboles, un mínimo de 3 instancias por nodo, y una proporción de submuestreo de columnas de 0.1 por árbol. En el caso de LightGBM, el mejor desempeño se logró con una tasa de aprendizaje también de 0.01, una profundidad máxima de 3, 800 árboles, una proporción de submuestreo de columnas de 0.1 y un número máximo de 8 hojas por árbol, lo que refleja una estructura más simple y menos profunda en comparación con XGBoost.

- El desempeño de los modelos desarrollados evidenció que XGBoost superó consistentemente a LightGBM en términos de precisión predictiva, mostrando menores valores de error (RMSE, MAE y MAPE) en la estimación de la resistencia a la compresión del concreto. Esta diferencia también se reflejó en las gráficas de dispersión, donde la relación entre valores predichos y experimentales fue más ajustada en el caso de XGBoost, confirmando su mayor capacidad para modelar la complejidad no lineal de los datos.
- Las pendientes de regresión superiores a 0.8 indicaron una buena concordancia global entre los valores experimentales y los predichos por ambos modelos, lo que sugiere una adecuada capacidad de generalización. No obstante, a pesar de que XGBoost presentó pendientes ligeramente menores en los conjuntos de entrenamiento y prueba en comparación con LightGBM, logró alcanzar un coeficiente de determinación (R^2) más alto, lo cual implica que el modelo explicó una mayor proporción de la variabilidad de los datos. Esta diferencia se puede atribuir a la arquitectura de XGBoost, que optimiza de manera más eficiente la función de pérdida mediante boosting secuencial regularizado, permitiéndole capturar relaciones no lineales más complejas con menor sobreajuste. En consecuencia, XGBoost mostró una mayor precisión y robustez tanto en la fase de entrenamiento como en la de validación, consolidándose como un modelo más confiable para la predicción de la resistencia a la compresión del concreto.

REFERENCIAS

- [1] O. Alshboul, G. Almasabha, A. Shehadeh, y K. Al-Shboul, «A comparative study of LightGBM, XGBoost, and GEP models in shear strength

management of SFRC-SBWS», Structures, vol. 61, p. 106009, mar. 2024, doi: 10.1016/j.istruc.2024.106009.

[2] A. Khashman y P. Akpinar, «Non-Destructive Prediction of Concrete Compressive Strength Using Neural Networks», Procedia Computer Science, vol. 108, pp. 2358-2362, ene. 2017, doi: 10.1016/j.procs.2017.05.039.

[3] K. C. Mouli et al., «Performance analysis of linear and non-linear machine learning models for forecasting compressive strength of concrete», Cogent Engineering, vol. 11, n.o 1, p. 2368101, dic. 2024, doi: 10.1080/23311916.2024.2368101.

[4] S. Pandey, S. Paudel, K. Devkota, K. Kshetri, y P. G. Asteris, «Machine learning unveils the complex nonlinearity of concrete materials' uniaxial compressive strength», International Journal of Construction Management, vol. 0, n.o 0, pp. 1-15, abr. 2024, doi: 10.1080/15623599.2024.2345008.

[5] Y. Pawar y Shrikant Rajendra Kate, «Curing of Concrete: A Review», 2020, doi: 10.13140/RG.2.2.32095.07848.

[6] V. Cominato, F. Benavente Canteras, L. Andréia Gachet, y R. Cristina Cecche Lintz, «The effect of granulometry of natural and recycled coarse aggregate on permeable concrete properties», Materials Today: Proceedings, vol. 65, pp. 1711-1718, ene. 2022, doi: 10.1016/j.matpr.2022.04.717.

[7] F. Cao et al., «Compressive strength prediction of Portland cement clinker using machine learning by XRD data», Journal of Sustainable Cement-Based Materials, vol. 13, n.o 8, pp. 1120-1131, ago. 2024, doi: 10.1080/21650373.2024.2351461.

[8] H. Constantinescu, O. Gherman, C. Negrutiu, y S. P. Ioan, «Mechanical Properties of Hardened High Strength Concrete», Procedia Technology, vol. 22, pp. 219-226, ene. 2016, doi: 10.1016/j.protcy.2016.01.047.

[9] B. W. Chong et al., «Design of Experiment on Concrete Mechanical Properties Prediction: A Critical Review», Materials, vol. 14, n.o 8, Art. n.o 8, ene. 2021, doi: 10.3390/ma14081866.

- [10] F. Khademi, M. Akbari, S. M. Jamal, y M. Nikoo, «Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete», *Front. Struct. Civ. Eng.*, vol. 11, n.o 1, pp. 90-99, mar. 2017, doi: 10.1007/s11709-016-0363-9.
- [11] J. L. Santamaría, B. Adame, C. Bermeo, J. L. Santamaría, B. Adame, y C. Bermeo, «Influencia de la calidad de los agregados y tipo de cemento en la resistencia a la compresión del hormigón dosificado al volumen», *Revista Digital Novasinergia*, vol. 4, n.o 1, pp. 91-101, may 2021, doi: 10.37135/ns.01.07.05.
- [12] M. Z. Naser, «Integrating Machine Learning Models into Building Codes and Standards: Establishing Equivalence through Engineering Intuition and Causal Logic», *Journal of Structural Engineering*, vol. 150, n.o 5, p. 04024039, may 2024, doi: 10.1061/JSENDH.STENG-12934.
- [13] A. H. A. Ahmed, W. Jin, y M. A. H. Ali, «Prediction of compressive strength of recycled concrete using gradient boosting models», *Ain Shams Engineering Journal*, vol. 15, n.o 9, p. 102975, sep. 2024, doi: 10.1016/j.asej.2024.102975.
- [14] P. C. Chiadighikaobi, M. Hematibahar, M. Kharun, N. A. Stashevskaya, y K. Camara, «Predicting mechanical properties of self-healing concrete with *Trichoderma Reesei* Fungus using machine learning», *Cogent Engineering*, vol. 11, n.o 1, p. 2307193, dic. 2024, doi: 10.1080/23311916.2024.2307193.
- [15] L. A. Bedriñana, J. Sucasaca, J. Tovar, y H. Burton, «Design-Oriented Machine-Learning Models for Predicting the Shear Strength of Prestressed Concrete Beams», *Journal of Bridge Engineering*, vol. 28, n.o 4, p. 04023009, abr. 2023, doi: 10.1061/JBENF2.BEENG-6013.
- [16] B. P. Koya, S. Aneja, R. Gupta, y C. Valeo, «Comparative analysis of different machine learning algorithms to predict mechanical properties of concrete», *Mechanics of Advanced Materials and Structures*, vol. 29, n.o 25, pp. 4032-4043, oct. 2022, doi: 10.1080/15376494.2021.1917021.
- [17] H. Wang, J. Tang, M. Wu, X. Wang, y T. Zhang, «Application of machine learning missing data imputation techniques in clinical decision making: taking

the discharge assessment of patients with spontaneous supratentorial intracerebral hemorrhage as an example», BMC Med Inform Decis Mak, vol. 22, n.o 1, p. 13, dic. 2022, doi: 10.1186/s12911-022-01752-6.

[18] P. Lu, S. Chen, y Y. Zheng, «Artificial Intelligence in Civil Engineering», Mathematical Problems in Engineering, vol. 2012, n.o 1, p. 145974, ene. 2012, doi: 10.1155/2012/145974.

[19] T. Han, A. Siddique, K. Khayat, J. Huang, y A. Kumar, «An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete», Construction and Building Materials, vol. 244, p. 118271, may 2020, doi: 10.1016/j.conbuildmat.2020.118271.

[20] S. Moon y A. Munira Chowdhury, «Utilization of Prior Information in Neural Network Training for Improving 28-Day Concrete Strength Prediction», Journal of Construction Engineering and Management, vol. 147, n.o 5, p. 04021028, may 2021, doi: 10.1061/(ASCE)CO.1943-7862.0002047.

[21] M. Mohtasham Moein et al., «Predictive models for concrete properties using machine learning and deep learning approaches: A review», Journal of Building Engineering, vol. 63, p. 105444, ene. 2023, doi: 10.1016/j.jobbe.2022.105444.

[22] Q. Wang, J. Qi, S. Hosseini, H. Rasekh, y J. Huang, «ICA-LightGBM Algorithm for Predicting Compressive Strength of Geo-Polymer Concrete», Buildings, vol. 13, n.o 9, Art. n.o 9, sep. 2023, doi: 10.3390/buildings13092278.

[23] H.-V. T. Mai, M. H. Nguyen, y H.-B. Ly, «Development of machine learning methods to predict the compressive strength of fiber-reinforced self-compacting concrete and sensitivity analysis», Construction and Building Materials, vol. 367, p. 130339, feb. 2023, doi: 10.1016/j.conbuildmat.2023.130339.

[24] S. Li y J. Y. Richard Liew, «Experimental and Data-Driven analysis on compressive strength of steel fibre reinforced high strength concrete and mortar at elevated temperature», Construction and Building Materials, vol. 341, p. 127845, jul. 2022, doi: 10.1016/j.conbuildmat.2022.127845.

- [25] X.-Y. Zhao, J.-X. Chen, y B. Wu, «An interpretable ensemble-learning-based open source model for evaluating the fire resistance of concrete-filled steel tubular columns», *Engineering Structures*, vol. 270, p. 114886, nov. 2022, doi: 10.1016/j.engstruct.2022.114886.
- [26] Z. Mei, J. Yu, W. Ding, L. Kong, y J. Zhao, «Bus Passenger Volume Forecasting Model Based on XGBoost Integrated Learning Algorithm», pp. 3100-3113, dic. 2020, doi: 10.1061/9780784483053.261.
- [27] A. Abdulalim Alabdullah, M. Iqbal, M. Zahid, K. Khan, M. Nasir Amin, y F. E. Jalal, «Prediction of rapid chloride penetration resistance of metakaolin based high strength concrete using light GBM and XGBoost models by incorporating SHAP analysis», *Construction and Building Materials*, vol. 345, p. 128296, ago. 2022, doi: 10.1016/j.conbuildmat.2022.128296.
- [28] M. Nithurshan y Y. Elakneswaran, «A systematic review and assessment of concrete strength prediction models», *Case Studies in Construction Materials*, vol. 18, p. e01830, jul. 2023, doi: 10.1016/j.cscm.2023.e01830.
- [29] Q.-F. Li y Z.-M. Song, «High-performance concrete strength prediction based on ensemble learning», *Construction and Building Materials*, vol. 324, p. 126694, mar. 2022, doi: 10.1016/j.conbuildmat.2022.126694.
- [30] J. Pal y D. Chakrabarty, «Effects of input/output parameters on artificial neural network model efficiency for breakthrough contaminant prediction», *Water Supply*, vol. 21, n.o 7, pp. 3614-3628, abr. 2021, doi: 10.2166/ws.2021.125.
- [31] S. Paudel, A. Pudasaini, R. K. Shrestha, y E. Kharel, «Compressive strength of concrete material using machine learning techniques», *Cleaner Engineering and Technology*, vol. 15, p. 100661, ago. 2023, doi: 10.1016/j.clet.2023.100661.
- [32] A. Behnood y E. M. Golafshani, «Machine learning study of the mechanical properties of concretes containing waste foundry sand», *Construction and Building Materials*, vol. 243, p. 118152, may 2020, doi: 10.1016/j.conbuildmat.2020.118152.

[33] C. Deepa, K. Sathiyakumari, y V. P. Sudha, «Prediction of the Compressive Strength of High Performance Concrete Mix using Tree Based Modeling», IJCA, vol. 6, n.o 5, pp. 18-24, sep. 2010, doi: 10.5120/1076-1406.

[34] D.-K. Bui, T. Nguyen, J.-S. Chou, H. Nguyen-Xuan, y T. D. Ngo, «A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete», Construction and Building Materials, vol. 180, pp. 320-333, ago. 2018, doi: 10.1016/j.conbuildmat.2018.05.201.

[35] L. Acuña Pinaud, P. C. Espinoza Haro, I. Moromi Nakata, A. V. Torre Carrillo, y F. García Fernández, «Concreto de alto rendimiento, predicción de su resistencia a la compresión mediante redes neuronales artificiales», TECNIA, vol. 27, n.o 1, Art. n.o 1, jun. 2017, doi: 10.21754/tecnia.v27i1.125.

[36] M. Noman, A. Khattak, Z. Alam, M. Yaqub, y E. Noroozinejad Farsangi, «Predicting the Residual Compressive Strength of Concrete Exposed to Elevated Temperatures Using Interpretable Machine Learning», Practice Periodical on Structural Design and Construction, vol. 29, n.o 4, p. 04024055, nov. 2024, doi: 10.1061/PPSCFX.SCENG-1536.

[37] X. Ying, «An Overview of Overfitting and its Solutions», J. Phys.: Conf. Ser., vol. 1168, p. 022022, feb. 2019, doi: 10.1088/1742-6596/1168/2/022022.

[38] J. H. Friedman, «Greedy function approximation: A gradient boosting machine.», Ann. Statist., vol. 29, n.o 5, oct. 2001, doi: 10.1214/aos/1013203451.

[39] A. Natekin y A. Knoll, «Gradient boosting machines, a tutorial», Front. Neurobot., vol. 7, 2013, doi: 10.3389/fnbot.2013.00021.

[40] G. Ke et al., «LightGBM: A Highly Efficient Gradient Boosting Decision Tree», en Advances in Neural Information Processing Systems, Curran Associates, Inc., 2017. Accedido: 6 de agosto de 2024. [En línea]. Disponible en: https://papers.nips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html

[41] T. Chen y C. Guestrin, «XGBoost: A Scalable Tree Boosting System», en Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA: ACM, ago. 2016, pp. 785-794. doi: 10.1145/2939672.2939785.

[42] M. I. Khan et al., «Effective use of recycled waste PET in cementitious grouts for developing sustainable semi-flexible pavement surfacing using artificial neural network (ANN)», Journal of Cleaner Production, vol. 340, p. 130840, mar. 2022, doi: 10.1016/j.jclepro.2022.130840.

[43] M. Iqbal, Q. Zhao, D. Zhang, F. E. Jalal, y A. Jamal, «Evaluation of tensile strength degradation of GFRP rebars in harsh alkaline conditions using non-linear genetic-based models», Mater Struct, vol. 54, n.o 5, p. 190, sep. 2021, doi: 10.1617/s11527-021-01783-x.

[44] M. Iqbal, D. Zhang, F. E. Jalal, y M. Faisal Javed, «Computational AI prediction models for residual tensile strength of GFRP bars aged in the alkaline concrete environment», Ocean Engineering, vol. 232, p. 109134, jul. 2021, doi: 10.1016/j.oceaneng.2021.109134.

Financiamiento de la investigación

Con recursos propios.

Declaración de intereses

Declaro no tener ningún conflicto de intereses, que puedan haber influido en los resultados obtenidos o las interpretaciones propuestas.

Declaración de consentimiento informado

El estudio se realizó respetando el Código de ética y buenas prácticas editoriales de publicación.

Derechos de uso

Copyright© 2025 por **Jaime Yelsin Rosales Malpartida, Victor Adrian Rosales-Cutipa**



[Este texto está protegido por la Licencia Creative Commons Atribución 4.0 Internacional.](https://creativecommons.org/licenses/by/4.0/)

Usted es libre para compartir, copiar y redistribuir el material en cualquier medio o formato y adaptar el documento, remezclar, transformar y crear a partir del material para cualquier propósito, incluso comercialmente, siempre que cumpla la condición de atribución: usted debe reconocer el crédito de una obra de manera adecuada, proporcionar un enlace a la licencia, e indicar si se han realizado cambios. Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que tiene el apoyo del licenciante o lo recibe por el uso que hace.