



Journal of Advanced Mining Modeling (JAMM)

Página web de la revista: <https://revistas.unitru.edu.pe/index.php/jamm/index>
Vol. 2, Nº 1, pp. 9-23, Enero – Junio 2026



Predicción de la resistencia a la compresión del hormigón mediante regresión lineal múltiple (OLS) y preprocesamiento de datos

Adrian Ruiz-Victorio^{1*} , Rosas Simón Cruz-Romero² , Orlando Alex Siccha-Ruiz³ 

¹. Departamento de Ingeniería de Minas / Facultad de Ingeniería / Universidad Nacional de Trujillo, Trujillo, Perú

². Departamento de Ingeniería de Minas / Facultad de Ingeniería / Universidad Nacional de Trujillo, Trujillo, Perú

³. Departamento de Ingeniería de Minas / Facultad de Ingeniería / Universidad Nacional de Trujillo, Trujillo, Perú

* Autor de correspondencia: h5053401521@unitru.edu.pe (A. Ruiz-Victorio)

Recibido 19 de diciembre de 2025

Revisado 29 de diciembre de 2025

Aceptado 01 de diciembre de 2026

Publicado en línea 01 de diciembre de 2026

Resumen

La seguridad y estabilidad de los sistemas de soporte y revestimiento constituyen un aspecto crítico en los proyectos de ingeniería minera, debido a las altas exigencias estructurales impuestas por las condiciones operativas, especialmente en minería subterránea. El objetivo principal de este estudio fue evaluar la capacidad de modelos de regresión lineal múltiple (OLS) para predecir la resistencia a la compresión del hormigón, analizando el efecto de diferentes técnicas de preprocesamiento de datos sobre su desempeño predictivo. La metodología adoptó un enfoque cuantitativo y se basó en el análisis de un conjunto de datos experimentales compuesto por más de 1000 mezclas de hormigón, considerando variables asociadas a la dosificación de los materiales y a la edad de curado. Se desarrollaron y compararon tres configuraciones de modelado: predictores sin transformación, predictores estandarizados y predictores transformados mediante la función $\log(x+1)$. El desempeño de los modelos se evaluó utilizando el Error Cuadrático Medio (MSE) y el coeficiente de determinación R^2 . Los resultados obtenidos muestran que los modelos construidos con datos originales y estandarizados presentaron un comportamiento similar, con valores de R^2 cercanos a 0.351 en el conjunto de prueba. En contraste, la transformación logarítmica permitió reducir el MSE de prueba a aproximadamente 55 MPa² e incrementar el R^2 hasta valores próximos a 0.750, evidenciando una mejora sustancial en la precisión y capacidad de generalización del modelo. El alcance de estos resultados se orienta al uso del modelo como una herramienta predictiva de apoyo de control de calidad del hormigón en aplicaciones mineras. En conclusión, la regresión lineal múltiple combinada con un preprocesamiento adecuado mediante transformación logarítmica constituye una alternativa eficiente y confiable para la estimación de la resistencia a la compresión del hormigón, manteniendo una interpretación clara de los resultados y aportando soporte práctico para la toma de decisiones técnicas en minería.

Palabras clave: Resistencia; Hormigón; Regresión lineal; Preprocesamiento de datos; Aprendizaje automático.

Abstract

The safety and stability of ground support and lining systems are critical aspects of mining engineering projects, due to the high structural demands imposed by operational conditions, particularly in underground mining. The main objective of this study was to evaluate the capability of multiple linear regression models (OLS) to predict concrete compressive strength, analyzing the effect of different data preprocessing techniques on their predictive performance. A quantitative approach was adopted, based on the analysis of an experimental dataset comprising more than 1,000 concrete mix designs, considering variables related to material proportions and curing age. Three modeling configurations were developed and compared: untransformed predictors, standardized predictors, and predictors transformed using the $\log(x+1)$ function. Model performance was assessed using Mean Squared Error (MSE) and the coefficient of determination (R^2). The results indicate that models built with original and standardized data exhibited similar behavior, with R^2 values close to 0.351 on the test dataset. In contrast, logarithmic transformation significantly improved model performance, reducing the test MSE to approximately 55 MPa² and increasing R^2 to values close to 0.750, demonstrating a substantial enhancement in accuracy

and generalization capability. These results support the use of the proposed model as a predictive tool for concrete quality control in mining applications. In conclusion, multiple linear regression combined with appropriate logarithmic preprocessing represents an efficient and reliable alternative for estimating concrete compressive strength, preserving model interpretability and providing practical support for technical decision-making in mining engineering.

Keywords: Compressive strength; Concrete; Linear regression; Data preprocessing; Machine learning.

1. Introducción

El hormigón es un material indispensable en la industria minera, valorado por su accesibilidad, costo razonable y adecuado desempeño estructural, especialmente para cimentaciones de plantas concentradoras, plataformas de lixiviación, rampas, taludes reforzados, entre otras áreas dentro de la minería. De acuerdo con los estudios técnicos, la resistencia a la compresión del hormigón es uno de los parámetros más relevantes en la ingeniería de minas, en estos contextos el hormigón y especialmente el hormigón proyectado, cumple una función en la seguridad operacional, ya que su desempeño mecánico influye directamente en la estabilidad del macizo rocoso y en la mitigación de riesgos geomecánicos [1]. En la minería subterránea, la variabilidad en la calidad del hormigón puede generar ciertas consecuencias significativas, puede abarcar desde fallas prematuras del sostenimiento hasta incrementos de costos operativos por sobreconsumos de materiales o reprocesos constructivos. Por ello, predecir la resistencia del hormigón de manera confiable, en función de la dosificación y las condiciones de curado, constituye un problema técnico de alto interés para ingenieros de minas y geotécnicos [2]. Tradicionalmente, este parámetro ha sido evaluado mediante ciertos ensayos experimentales, en los cuales, si bien son precisos, demandan tiempo, recursos y una planificación que no es siempre compatible con la dinámica operativa de una mina productiva [3]. Diversos estudios han demostrado que la resistencia en cuanto al hormigón depende de muchas variables interrelacionadas, tales como la presencia del cemento, así como la relación de agua & aglomerante, por otra parte, se incluye el uso de aditivos minerales como los aditivos químicos y la edad de curado. En aplicaciones mineras, estas variables suelen presentarse una mayor dispersión debido a restricciones logísticas, condiciones ambientales subterráneas y variaciones en los métodos de colocación del material, lo que incrementa la incertidumbre asociada al desempeño mecánico del hormigón [4]. A partir de estas consideraciones, el empleo de modelos estadísticos y técnicas de aprendizaje automático se ha consolidado como una alternativa viable para estimar la resistencia a la compresión a partir de registros históricos de mezclas. Diversos trabajos previos han demostrado que los modelos de regresión son capaces de identificar relaciones relevantes entre los componentes del hormigón y su resistencia final, siempre que el conjunto de datos sea adecuadamente preparado y analizado [5]. Dentro de este grupo, la regresión lineal, destaca por su simplicidad conceptual, así como por su facilidad de implementación en aplicaciones de ingeniería, especialmente en entornos industriales donde se requiere transparencia en los resultados obtenidos [6]. En el contexto minero, la interpretabilidad de los modelos predictivos adquiere una importancia particular, ya que permite a los ingenieros comprender de manera directa la influencia de cada variable sobre el comportamiento del material y respaldar decisiones técnicas relacionadas sobre la dosificación y el control de calidad. A diferencia de los modelos complejos considerados como “caja negra” los enfoques lineales lo hacen más fácil la validación técnica y en su incorporación en procedimientos operativos estándar propios de proyectos mineros [7]. Sin embargo, el empleo directo de modelos lineales sin tratamiento previo suele presentarse limitaciones que están asociadas a problemas estadísticos, en los cuales pueden afectar la capacidad de generalización del modelo y conducir a condiciones poco confiables de la resistencia del hormigón. Estos problemas pueden afectar de manera negativa la capacidad de generalización del modelo y conducir a estimaciones poco confiables de la resistencia [8]. Frente a esta situación, se ha propuesto el uso de técnicas de preprocesamiento de datos, como la estandarización y las transformaciones logarítmicas, con el fin de mejorar el desempeño que se han propuesto para mejorar los modelos lineales [9]. Investigaciones enfocadas en el diseño de mezclas de hormigón, han evidenciado que la transformación logarítmica contribuye a reducir la dispersión de datos, estabilizar la varianza y poner en evidencia las relaciones que no resultan claras en la escala original de los datos [10]. Este enfoque es especialmente en minería, donde las condiciones reales de operación difícilmente cumplen supuestos ideales y donde la robustez del modelo frente a conjuntos de datos no balanceados son resultados claves para su aplicación práctica [11]. En la minería subterránea, el hormigón proyectado cumple un rol fundamental como elemento de sostenimiento inmediato, por lo que su desempeño no solo se evalúa en función de resistencias finales, sino también mediante criterios de resistencia temprana establecidos en los estándares operáticos de

mina. En la práctica, se consideran valores mínimos del orden 2 MPa a las 2 horas y resistencias finales superiores a 28 MPa a los 28 días, como requisitos habituales para garantizar la estabilidad inicial y a largo plazo de las excavaciones. En este sentido, el desarrollo de herramientas predictivas basadas en datos históricos de dosificación permite anticipar el cumplimiento de dichos criterios, optimizar el control de calidad y reducir la dependencia exclusiva de ensayos destructivos, aportando una ventaja operativa directa en ambientes subterráneos donde el tiempo y la seguridad son viables críticas. Este trabajo, tiene como objetivo evaluar el desempeño y la capacidad de los modelos clásicos de regresión lineal para la predicción de la resistencia a la compresión del hormigón empleando en aplicaciones vinculadas a la ingeniería de minas, comparando de manera sistemática el efecto de distintos métodos de preprocesamiento de datos. A partir de un conjunto de datos ampliamente utilizado y validado en la literatura, se busca identificar qué transformaciones permiten mejorar la precisión, la capacidad de generalización y la interpretabilidad del modelo, con el propósito de ofrecer una herramienta analítica práctica para el diseño, y el control de mezclas de hormigón en operaciones mineras contemporáneas [12].

2. Materiales y métodos

El desarrollo del presente estudio, se sustenta en un conjunto de datos experimentales reales compuesto por más de 1000 mezclas de hormigón como diseñado para analizar y predecir la resistencia a la compresión de material en contextos técnicos afines a la ingeniería de minas, tales como sostenimiento subterráneo o superficial, revestimiento de túneles y estructuras auxiliares mineras cada registro del conjunto de datos representa una mezcla de hormigón caracterizada por su dosificación y edad de curado, junto con el valor medido de resistencia a la compresión en megapascales (MPa), propiedad mecánica fundamental para evaluar la estabilidad y seguridad estructural en excavaciones mineras [13]. Las variables incluidas reflejan componentes comúnmente utilizados en mezclas de hormigón empleando en operaciones mineras, donde las condiciones de colocación, curado y el control de calidad presentan una mayor variabilidad respecto a entornos convencionales [14].

2.1 Datos y variables

2.1.1 Variables del conjunto de datos experimental

En la Tabla 1, se muestran una descripción de las variables predictoras y la variable objetivo consideradas en el estudio.

Tabla 1. Variables del conjunto de datos experimental

Variable	Descripción	Unidad
Cement	Contenido de cemento en la mezcla	Kg/m ³
Slag	Escoria de alto horno	Kg/m ³
FlyAsh	Ceniza volante	Kg/m ³
Water	Contenido de agua	Kg/m ³
Superplasticizer	Aditivo superplastificante	Kg/m ³
CoarseAgg	Agregado grueso	Kg/m ³
FineAgg	Agregado fino	Kg/m ³
Age	Edad de curado	Días
Strength	Resistencia a la compresión del hormigón	MPa

2.2. Preparación y partición de datos

Previamente al modelo, los datos fueron sometidos a un proceso de verificación numérica con la finalidad de asegurar la coherencia de las magnitudes y la consistencia de variables. Posteriormente, el conjunto de agrupación de datos fue dividido en partes o subconjuntos, siendo el de entrenamiento, así como la prueba, siguiendo un esquema fijo que permite evaluar el desempeño del modelo sobre observaciones no utilizadas durante el ajuste [15].

En este procedimiento reproduce un escenario realista, donde los modelos predictivos deben mantener su capacidad de generalización frente a nuevas condiciones de mezcla y variaciones operativas propias del entorno minero [16].

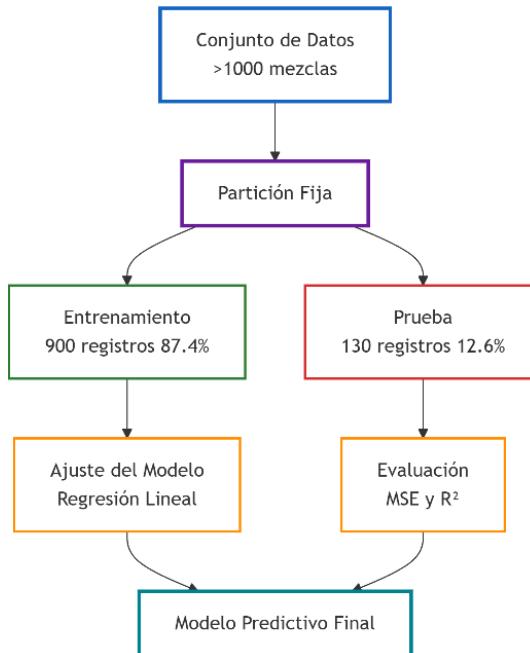


Figura 1. Esquema del proceso de partición de datos

2.3 Preprocesamiento de datos

Se tomó como referencia que el objetivo de identificar ciertos impactos de preprocesamiento en la capacidad predictiva del modelo, se evaluaron 3 configuraciones distintas de los datos de entrada, como predictores sin transformación, predictores estandarizados y predictores transformados mediante $\log(x+1)$.

La estandarización se realizó mediante siguiente la expresión:

$$Z_i = \frac{X_i - \mu}{\sigma} \quad (1)$$

Donde:

Z_i – estandarización

X_i – valor original de la variable,

μ – media de dicha variable,

σ – desviación estándar.

2.3.1 Transformación logarítmica aplicada a los predictores

Esta transformación se emplea para reducir la asimetría de los datos y estabilizar la varianza, mejorando el desempeño de modelos lineales en conjuntos de datos experimentales reales [18].

$$X'_i = \log(X_i + 1) \quad (2)$$

Donde:

X'_i – transformación logarítmica,

\log – logaritmo.

2.4 Modelado de la regresión lineal múltiple (OLS)

La estimación de acuerdo a la resistencia a la compresión del hormigón se ejecutó a través de un modelo de regresión lineal multivariado, ajustado utilizando un algoritmo de MCO (mínimos cuadrados ordinarios) [19]. El modelo general adoptado se expresa como:

$$y = \beta_0 + \sum_{i=1}^n \beta_i X_i + \varepsilon \quad (3)$$

Donde:

y – resistencia la compresión del hormigón (MPa),

β_0 – término independiente,

β_i – coeficientes de regresión,

X_i – variables predictoras,

ε – término del error.

En este enfoque fue seleccionada por su elevada interpretabilidad, característica esencial en ingeniería de minas, ya que permite identificar de forma directa la influencia de cada componente del hormigón sobre su desempeño mecánico y apoyar la toma de decisiones en control de calidad y diseño de mezclas [20].

2.5 Métricas de evaluación

Estas métricas permiten cuantificar la precisión del modelo y su capacidad de generalización, aspectos críticos para aplicaciones mineras donde errores de estimación pueden comprometer la seguridad estructural y la eficiencia operativa [21]. El desempeño de los modelos fue evaluado utilizando el MSE y R^2 en donde se mencionan posteriormente.

2.5.1 Error Cuadrático Medio (MSE)

Este modelo hace referencia a unos errores que se utiliza para cuantificar cierta precisión global en cada modelo predictivo. Por otros términos, el MSE permite evaluar que tan lejos se encuentra las reducciones del modelo respecto a los valores reales de resistencia del hormigón, lo cual es muy crucial considerando que una subestimación puede comprometer la seguridad del sostenimiento y una sobreestimación puede generar costos innecesarios en operaciones mientras.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2 \quad (4)$$

Donde:

MSE – error cuadrático medio,

n – número total de observaciones evaluadas,

y_i – resistencia real medida del hormigón,

y'_i – resistencia predicha por el modelo,

$(y_i - y'_i)^2$ – error cuadrático individual

2.5.2 Coeficiente de determinación R^2

Este R^2 es una métrica clave que cuantifica cierto poder explicativo del modelo en proporción de la variabilidad de la resistencia que puede ser explicada por el conjunto de variables predictoras. Estudios realizados, muestran que valor cercano a 1 significa que el modelo logra explicar gran parte del comportamiento mecánica del hormigón, algo fundamental cuando se evalúa la confiabilidad estructural del material en labores mineras donde el soporte debe responder adecuadamente a esfuerzo geomecánicos.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y_i - y''_i)^2} \quad (5)$$

Donde:

R^2 – coeficiente de determinación,

y_i – resistencia real medida del hormigón,

y'_i – resistencia predicha por el modelo,

y''_i – valor promedio de la resistencia real,

$(y_i - y'_i)^2$ – suma de errores al cuadrado,

$(y_i - y''_i)^2$ – variabilidad total presente en los datos.

3. Resultados

3.1. Desempeño predictivo de los modelos

El rendimiento de los tres modelos evaluados, los predictores sin transformación, los predictores estandarizados y los predictores transformados mediante $\log(x+1)$, se presenta en la Tabla 2, que muestra los valores del Error Cuadrático Medio (MSE) y (R^2) para los datos de entrenamiento y test para todos los predictores.

Tabla 2. Desempeño de los modelos en términos de MSE y R^2

Configuración del modelo	Train MSE	Test MSE	Train R^2	Test R^2
Predictores sin transformar	105.20	142.67	0.624	0.351
Predictores estandarizados	105.20	142.67	0.624	0.351
Predictores Log(x+1)	56.39	55.06	0.799	0.750

Los resultados muestran que los modelos basados en predictores sin transformación y predictores estandarizados presentan un desempeño prácticamente idéntico. En ambos casos, el error de predicción en el conjunto de prueba es elevado y el valor de R^2 indica una capacidad limitada para explicar la variabilidad de la resistencia a la compresión.

En contraste, el modelo que emplea predictores transformados mediante $\log(x+1)$ muestra una reducción sustancial del error, con valores de MSE cercanos a la mitad de los obtenidos por los otros enfoques. Asimismo, el incremento del R^2 en el conjunto de prueba evidencia una mejora significativa en la capacidad de generalización del modelo.

3.2. Comparación grafica del error cuadrático medio

La Figura 2 presenta la comparación visual del MSE para los tres modelos, tanto en entrenamiento como en prueba. En esta figura se observa claramente que los modelos con predictores sin transformar y estandarizados presentan errores elevados y una diferencia marcada entre entrenamiento y prueba, lo que sugiere un ajuste poco robusto.

Por el contrario, el modelo con transformación logarítmica muestra valores de MSE similares en ambos conjuntos, lo que indica un comportamiento más estable y consistente frente a datos no utilizados durante el ajuste.

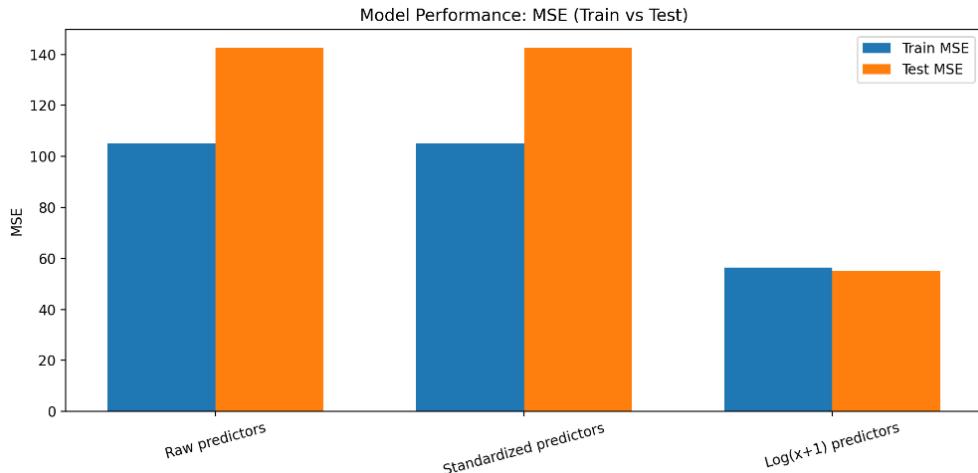


Figura 2. Comparación del Error Cuadrático Medio (MSE) para los tres modelos evaluados

3.3. Análisis del coeficiente de determinación (R^2)

La Figura 3 presenta una comparación del coeficiente de determinación R^2 para los tres métodos de modelado. Los modelos sin transformación y estandarizados explican solo una parte limitada de la variabilidad en la resistencia a la compresión en el conjunto de prueba, con valores alrededor de 0.351. Por otro lado, el modelo que utiliza predictores transformados mediante $\log(x+1)$ consigue explicar aproximadamente el 75% de la variabilidad observada en los datos de prueba, lo que refleja un ajuste notablemente superior y una mejor representación de las relaciones entre las variables de entrada y la resistencia a la compresión.

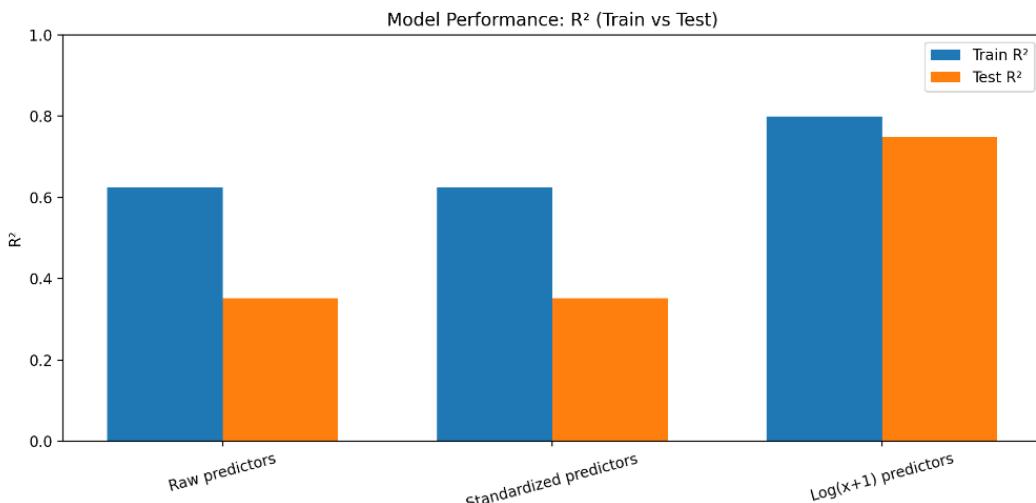


Figura 3. Comparación del coeficiente de determinación (R^2) en entrenamiento y prueba

3.4. Significancia estadística

La significancia estadística de las variables predictoras fue evaluada a partir de los modelos ajustados mediante mínimos cuadrados ordinarios. Para cada configuración de preprocessamiento se generaron tablas de valores p , correspondientes a los modelos con predictores sin transformar, estandarizados y transformados mediante $\log(x+1)$.

La Figura 4 muestra el gráfico de $-\log_{10}(p\text{-value})$ correspondiente al modelo con predictores estandarizados, donde se observa una clara diferencia en la relevancia estadística de las variables. De manera análoga, los resultados gráficos correspondientes a los modelos sin transformación y con transformación logarítmica se muestran en las Figuras 5 y 6, respectivamente.

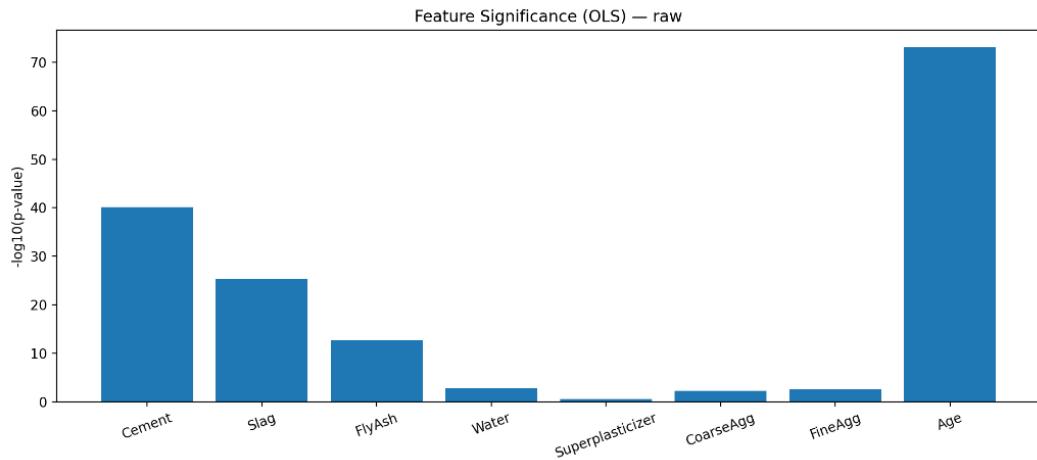


Figura 4. Significancia estadística de las variables predictoras en el modelo que utiliza predictores sin transformar, expresado como $-\log_{10}(p\text{-value})$

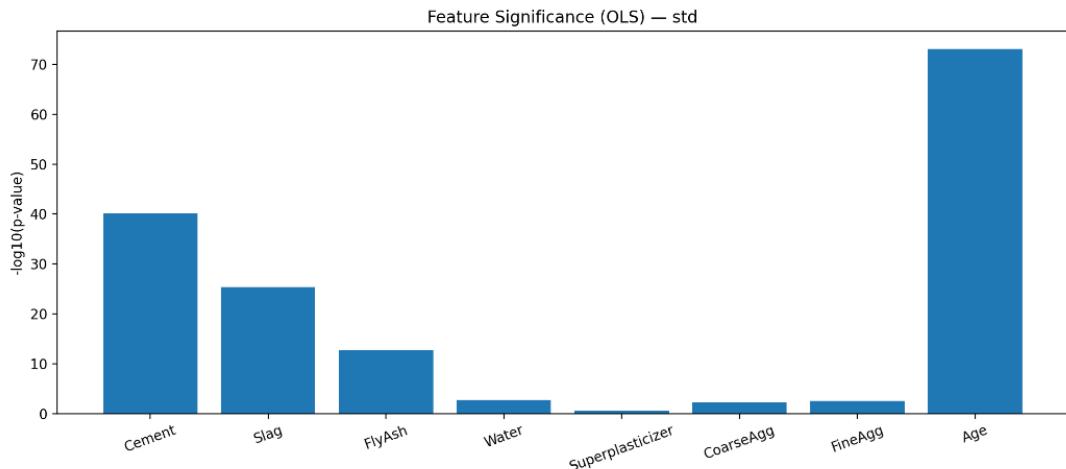


Figura 5. Significancia estadística de las variables predictoras en el modelo que utiliza predictores estandarizados, expresado como $-\log_{10}(p\text{-value})$

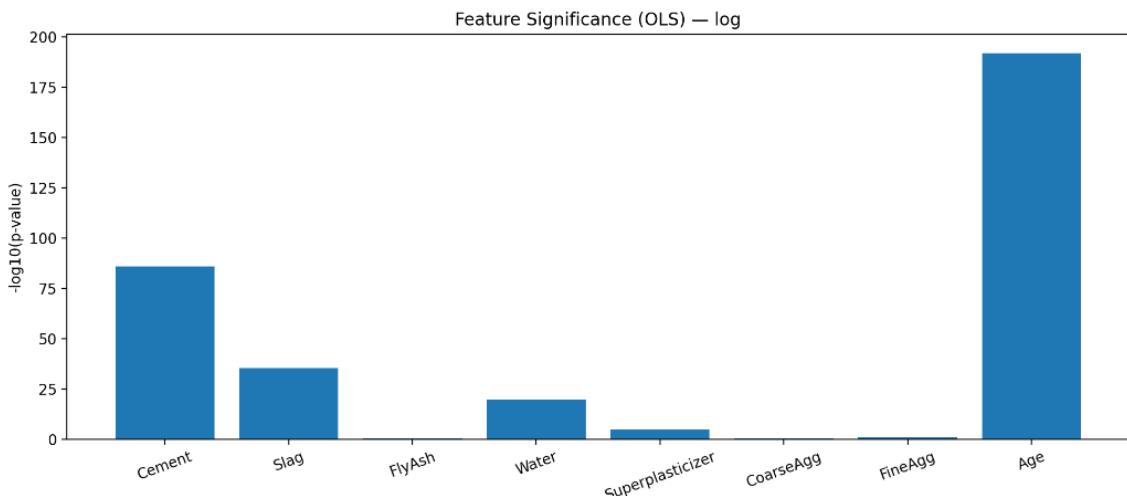


Figura 6. Significancia estadística de las variables predictoras en el modelo con transformación $\log(x+1)$, expresada como $-\log_{10}(p\text{-value})$

Los valores p obtenidos para cada modelo se presentan de forma detallada en las Tablas 3, 4 y 5.

Tabla 3. Valores p del modelo con predictores sin transformar

Predictor	p-value
Const	7.082e-02
Cement	6.971e-41
Slag	4.145e-26
FlyAsh	1.895e-13
Water	1.918e-03
Superplasticizer	2.461e-01
CoarseAgg	5.443e-03
FineAgg	2.738e-03
Age	9.516e-74

Tabla 4. Valores p del modelo con predictores estandarizados

Predictor	p-value
Const	0.000e+00
Cement	6.971e-41
Slag	4.145e-26
FlyAsh	1.895e-13
Water	1.918e-03
Superplasticizer	2.461e-01
CoarseAgg	5.443e-03
FineAgg	2.738e-03
Age	9.516e-74

Tabla 5. Valores p del modelo con predictores transformados mediante $\log(x+1)$

Predictor	p-value
Const	2.930e-01
Cement	1.289e-86
Slag	4.634e-36
FlyAsh	2.632e-01
Water	1.837e-20
Superplasticizer	1.021e-05
CoarseAgg	4.085e-01
FineAgg	9.428e-02
Age	1.710e-192

El análisis de estas Tablas indica que, en el modelo que utiliza la transformación $\log(x+1)$, el contenido de cemento y la edad de curado tienen los valores p más bajos. Esto sugiere que son muy significativos en la predicción de la compresión. También se encuentra que el contenido de agua es estadísticamente importante, aunque tiene un efecto negativo constante sobre la resistencia. Después de la transformación logarítmica, el superplasticificante se vuelve más relevante desde el punto de vista estadístico, mientras que la ceniza volante y los agregados muestran valores p altos, lo que sugiere que su impacto es limitado en el modelo lineal evaluado.

3.5. Validación y aplicación predictiva del modelo

Se validaron tres modelos de regresión, tales como, OLS sin transformar, estandarizado y con transformación logarítmica, utilizando un conjunto de datos independiente. Los modelos sin transformar y estandarizado mostraron un rendimiento limitado $R^2=0.351$ y patrones de residuos no aleatorios, indicando heterocedasticidad. En contraste, el modelo con transformación logarítmica $\log(x+1)$ demostró una capacidad predictiva superior $R^2=0.750$ y residuos bien comportados, validándolo como la herramienta más robusta, observándose en la Figura 7.

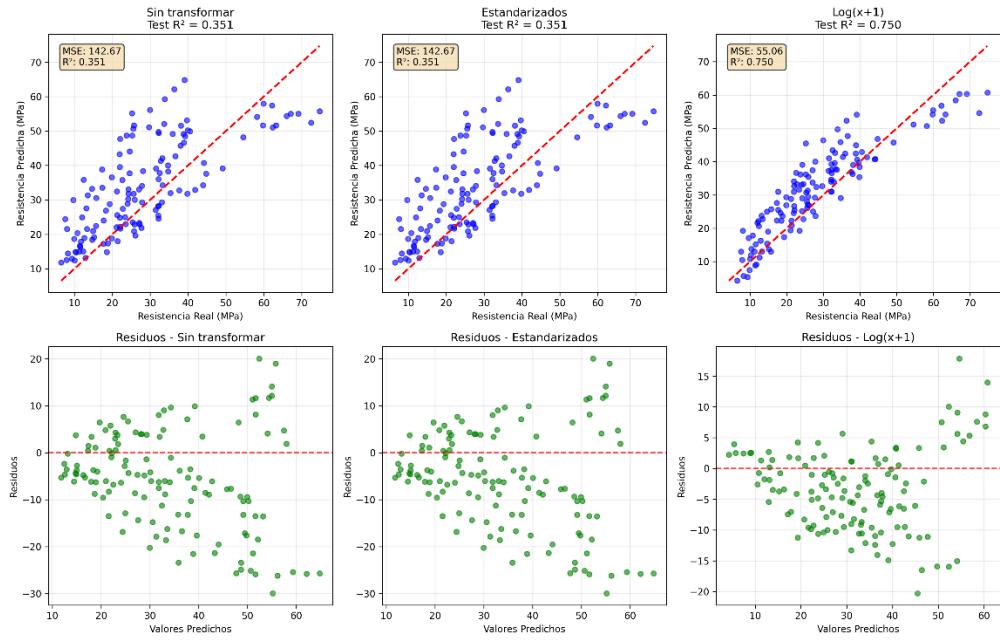
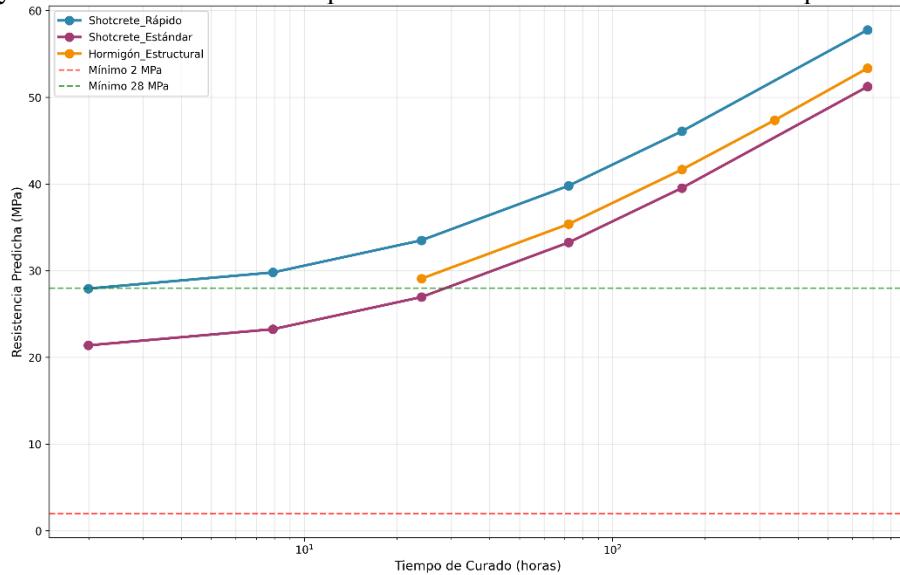


Figura 7. Validación de modelos OLS

No obstante, para evaluar su utilidad operativa, el modelo seleccionado se aplicó a tres escenarios de dosificación simulados: mezcla estándar, alto contenido de cemento y mezcla optimizada con aditivos. La Figura 8 compara las curvas de resistencia predichas con los estándares críticos de la mina: resistencia temprana (ej. 2 MPa a 2 horas) y resistencia final de diseño (ej. >28 MPa a 28 días). La herramienta permite predecir visual y cuantitativamente el cumplimiento de estos criterios antes de la aplicación en campo.

Figura 8. Curvas de desarrollo con transformación logarítmica $\log(x+1)$

Estos resultados confirman que el modelo logarítmico no solo es estadísticamente sólido, sino también una herramienta predictiva práctica para optimizar dosificaciones, garantizar el cumplimiento de especificaciones técnicas y reducir la incertidumbre en la fortificación subterránea.

4. Discusión

Los resultados obtenidos muestran que el tratamiento previo de los datos juega un papel crucial en la eficacia de los modelos de regresión lineal que se utilizan para la predicción de la resistencia a la compresión del hormigón. En concreto, el modelo que emplea predictores transformados mediante $\log(x+1)$ logró un R^2 de 0.80 en el conjunto de entrenamiento y de 0.75 en el conjunto de prueba, logrando un desempeño muy superior

al de los modelos que utilizan predictores sin transformar y estandarizados, los cuales presentaron un R^2 en pruebas de solo alrededor de 0.35. Esta marcada mejora coincide con lo indicado por Yeh [22], quien evidencia que la no linealidad natural de las propiedades mecánicas del hormigón limita el rendimiento de los modelos lineales cuando se utilizan datos en su escala original.

Desde la perspectiva del error en las predicciones, el modelo que utilizó una transformación logarítmica, logró reducir el Error Cuadrático Medio (MSE) en el conjunto de prueba, pasando de cifras superiores a 140 MPa² (modelos raw y estandarizados) aproximadamente de 55 MPa², lo que significa una reducción cercana al 60 %. Este descenso es especialmente importante en el campo de la ingeniería, dado que una estimación incorrecta de la resistencia del hormigón puede resultar en decisiones innecesariamente cautelosas o, en el peor escenario, generar riesgos estructurales [23].

El hecho de que los modelos con predictores sin transformar y estandarizados presenten métricas prácticamente idénticas (Train MSE \approx 105 MPa² y Test MSE \approx 143 MPa²) indica que la estandarización no corrige los problemas de asimetría y heterocedasticidad presentes en el conjunto de datos experimental. Este comportamiento ha sido observado previamente por Chou et al. [24], quienes señalan que, en propiedades mecánicas del hormigón, las transformaciones de tipo logarítmico resultan más efectivas que la normalización clásica cuando se emplean modelos de regresión lineal.

Dentro del esquema que usa la transformación $\log(x+1)$, la cantidad de cemento y cuánto tiempo se dejó curar tuvieron los valores p más pequeños, señalando que influyen mucho y de manera notable en la resistencia a la compresión. Este descubrimiento encaja con estudios anteriores que señalan estos dos elementos como los que más dirigen cómo el hormigón gana resistencia, ya sea en construcciones comunes o en el hormigón lanzado que se usa para asegurar túneles mineros [25]. Desde un punto de vista práctico, esto confirma que es vital manejar con mucho cuidado la cantidad de cemento y los tiempos de fraguado.

El nivel de agua tuvo un efecto importante, pero contrario, un resultado que coincide con la idea tradicional de la proporción agua a aglomerante. En la minería subterránea, donde la humedad del ambiente y cómo se aplica el hormigón pueden causar cambios en el agua real usada, este descubrimiento es particularmente importante. Neville [26], subraya que hasta aumentos chicos en el agua pueden causar bajas notables en la resistencia, algo que se ve bien en los números que el modelo calculó.

Un dato importante es que el superplastificante alcanzó significancia estadística después de aplicar la transformación logarítmica, lo que da a entender que existe una conexión curva entre este aditivo y la resistencia a la compresión. Este patrón se ha visto en investigaciones sobre hormigón proyectado, donde el superplastificante ayuda a que el material sea más fácil de trabajar y se compacte mejor, sin bajar la resistencia si se usa en las cantidades correctas [27]. Por otro lado, la ceniza volante y los agregados mostraron valores p altos, señalando que su aporte fue pequeño dentro del modelo lineal, lo que podría significar que su impacto se ve en interacciones que este método no logra captar.

El hecho de que los valores de MSE sean parecidos en el entrenamiento y la prueba para el modelo $\log(x+1)$ (casi 56 MPa² y cerca de 55 MPa², en ese orden) demuestra que se puede generalizar bien, algo fundamental para usarse en labores mineras. Tal como mencionan Asteris et al. [28], un modelo que se mantenga firme con datos nuevos es más de fiar que uno que se ajuste mucho al entrenar, pero rinda mal en situaciones inéditas, sobre todo donde los materiales cambian mucho.

Visto de modo práctico, un modelo de regresión lineal funciona bien. Este modelo puede emplearse como una herramienta de apoyo para el control de calidad del hormigón. Esto pasa en trabajos de minería, el costo de usarlo es bajo. Además, es fácil de entender, el cual ayuda a ponerlo en sistemas de control y se puede estimar pronto si una alcanzará la resistencia requerida. Pero, los resultados también sugieren algo, donde para ver efectos complejos entre componentes como agregados y adiciones minerales, se deben evaluara otros modelos no lineales en estudios futuros. Los resultados obtenidos adquieren relevancia al contrastarse con los criterios reales empleados en minas subterráneas para el control del hormigón. La capacidad del modelo con transformación $\log(x+1)$ para estimar la resistencia a la compresión con errores reducidos permite evaluar, de manera anticipada, si una mezcla determinada tiene alta probabilidad de alcanzar resistencias tempranas del orden de 2 MPa en las primeras horas de curado, así como resistencias finales superiores a 28 MPa a los 28 días. Esta comparación con estándares de comparación real evidencia el potencial del modelo como una herramienta predictiva de apoyo al control de calidad, orienta a la toma de decisiones en campo, más que como un reemplazo directo de los ensayos normativos.

Finalmente, nos obstante, los resultados logrados se mantienen estables y firmes dentro de la información examinada, usar este modelo en diferentes entornos geológicos necesitará cambios y ajustes de vez en cuando.

Ahmad et al. [30] indican que las variaciones en el origen de los materiales y en cómo se fraguan pueden influir en la certeza de los modelos que predicen, así que ajustar constantemente es fundamental para conservar la fiabilidad. De manera global, lo que encontramos demuestra que juntar datos de pruebas reales, preprocesamiento adecuado y usar modelos estadísticos sencillos de entender, es un método bueno para optimizar el manejo del hormigón en el sector de la minería

5. Conclusiones

En conclusión, los hallazgos muestran la importancia de la preparación de datos para los modelos de regresión lineal utilizados para pronosticar la resistencia a la compresión del desempeño del hormigón. Específicamente, la transformación $\log(x+1)$ aumentó considerablemente la capacidad explicativa del modelo, elevando el coeficiente de determinación en el conjunto de prueba de aproximadamente 0.351 a prácticamente 0.750.

Usando los predictores $\log(x+1)$, el modelo disminuyó el Error Cuadrático Medio en el conjunto de prueba en alrededor del 60% en comparación con los modelos estandarizados sin modificar, reduciéndolo así de más de 140 MPa² a aproximadamente 55 MPa². Esta mejora es particularmente crucial ya que la estabilidad y la seguridad dependen de una medición exacta de la resistencia del hormigón.

El análisis de significancia estadística reveló que las variables que más influyen en la resistencia a la compresión del hormigón son la cantidad de cemento y el tiempo de curado. Por el contrario, la cantidad de agua tuvo un impacto negativo constante. Estos resultados respaldan las técnicas establecidas de diseño de mezclas y llaman la atención sobre la necesidad de un control estricto de estos elementos en el hormigón empleado como revestimiento y soporte en operaciones mineras.

La significancia estadística del superplastificante aumenta tras la transformación logarítmica, lo que sugiere que su influencia en la resistencia a la compresión no es lineal. Este hallazgo subraya la importancia de utilizar métodos de preprocesamiento adecuados.

Al investigar aditivos químicos en hormigón destinado a la minería, donde es fundamental mantener un equilibrio entre trabajabilidad y rendimiento mecánico, la similitud entre los errores de entrenamiento y de prueba obtenidos por el modelo $\log(x+1)$ muestra una buena generalización, lo que sugiere que el modelo puede utilizarse de forma fiable en mezclas reales frescas de hormigón dentro de rangos operativos comparables. Para su posible incorporación como herramienta de apoyo al control de calidad en operaciones mineras subterráneas, esta característica es esencial.

No obstante, aunque el desempeño del modelo de regresión lineal con transformación $\log(x+1)$ fue claramente superior, con valores de R^2 cercanos a 0.750 y errores de predicción considerablemente reducidos, su precisión continúa estando fuertemente condicionada por la calidad y representatividad de los datos experimentales disponibles. Este modelo puede emplearse como una herramienta predictiva complementaria para verificar en etapas la probabilidad de cumplimiento de los criterios operativos de resistencia del hormigón, tales como resistencias iniciales cercanas a 2 MPa en las primeras horas y resistencias finales mayores a 28 MPa a los 28 días. Su aplicación práctica permitirá optimizar los programas de control de calidad, reducir reprocesos y fortalecer la gestión preventiva de la seguridad geomecánica.

Es determinante identificar que la presencia de variabilidad en los materiales, valores atípicos asociados a ensayos de laboratorio y posibles cambios en las condiciones de mezclado y curado del hormigón pueden afectar la estabilidad del modelo. En este sentido, resulta pertinente ampliar la base de datos mediante nuevas campañas experimentales, incorporar variables adicionales relacionadas con las condiciones ambientales y operativas propias de la minería subterránea, así como evaluar enfoques híbridos que integren modelos estadísticos con principios físico mecánicos del comportamiento del hormigón. A pesar de estas limitaciones, los resultados obtenidos constituyen un aporte relevante y validan el potencial de los métodos de análisis de datos y aprendizaje automático como herramientas de apoyo para mejorar la eficiencia, confiabilidad y predictibilidad del uso del hormigón en aplicaciones en minería u otra industria.

Futuros trabajos podrían además evaluar modelos híbridos que combinen enfoques estadísticos con principios físico-mecánicos del comportamiento del hormigón, con el objetivo de aumentar la robustez del modelo y su aplicabilidad en diferentes áreas.

6. Conflicto de Intereses

Los autores declaran que no tienen conflictos de interés.

7. Referencias

- [1]. Yeh, I.C. (1998). Modeling of strength of high-performance concrete using artificial neural networks. *Cement and Concrete Research*, 28(12), 1797–1808.
DOI: [https://doi.org/10.1016/S0008-8846\(98\)00165-3](https://doi.org/10.1016/S0008-8846(98)00165-3)
- [2]. Yeh, I.C. (2007). Modeling concrete strength with augment-neuron networks. *Cement and Concrete Research*, 37(7), 1063–1070.
DOI: <https://doi.org/10.1016/j.cemconres.2007.03.013>
- [3]. Asteris, P.G.; Mokos, V.G. (2020). Concrete compressive strength prediction using machine learning techniques. *Neural Computing and Applications*, 32, 15807–15828.
DOI: <https://doi.org/10.1007/s00521-019-04663-2>
- [4]. Chou, J.S.; Pham, A.D. (2013). Enhanced artificial intelligence for predicting high-performance concrete compressive strength. *Construction and Building Materials*, 49, 554–563.
DOI: <https://doi.org/10.1016/j.conbuildmat.2013.08.078>
- [5]. Neville, A.M. (2011). *Properties of Concrete* (5th ed.). Pearson Education Limited.
- [6]. Ahmad, A.; Farooq, F.; Ostrowski, K.A. (2021). Prediction of concrete compressive strength using machine learning algorithms. *Materials*, 14(15), 4141.
DOI: <https://doi.org/10.3390/ma14154141>
- [7]. Asteris, P.G.; Argyropoulos, I.; Cavalieri, L. (2021). Predicting concrete strength using hybrid machine learning models. *Applied Sciences*, 11(7), 3036.
DOI: <https://doi.org/10.3390/app11073036>
- [8]. Behnood, A.; Golafshani, E.M. (2018). Predicting the compressive strength of silica fume concrete using hybrid models. *Construction and Building Materials*, 170, 726–736.
DOI: <https://doi.org/10.1016/j.conbuildmat.2018.03.086>
- [9]. Ashrafian, A.; Nikoo, M.; Kadivar, A. (2020). Prediction of concrete strength using statistical learning methods. *Structures*, 28, 288–299.
DOI: <https://doi.org/10.1016/j.istruc.2020.08.021>
- [10]. Mansouri, I.; Hu, J.W. (2021). Application of machine learning techniques for predicting concrete strength. *Journal of Building Engineering*, 43, 102494.
DOI: <https://doi.org/10.1016/j.jobe.2021.102494>
- [11]. UCI Machine Learning Repository. (2019). Concrete Compressive Strength Dataset. University of California, Irvine.
DOI: <https://doi.org/10.7717/peerj-cs.154/table-6>
- [12]. Cortes, C.; Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297.
DOI: <https://doi.org/10.1007/BF00994018>
- [13]. Montgomery, D.C.; Peck, E.A.; Vining, G.G. (2012). *Introduction to Linear Regression Analysis* (5th ed.). Wiley.
DOI: <https://doi.org/10.1111/biom.12129>
- [14]. Kutner, M.H.; Nachtsheim, C.J.; Neter, J. (2005). *Applied Linear Statistical Models*. McGraw-Hill.
DOI: <https://doi.org/10.2307/1271154>

- [15]. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. (2021). An Introduction to Statistical Learning (2nd ed.). Springer.
DOI: <https://doi.org/10.32614/cran.package.islr2>
- [16]. Bishop, C.M. (2006). Pattern Recognition and Machine Learning. Springer.
DOI: https://doi.org/10.1007/978-0-387-45528-0_7
- [17]. Kuhn, M.; Johnson, K. (2013). Applied Predictive Modeling. Springer.
DOI: <https://doi.org/10.1007/978-1-4614-6849-3>
- [18]. Hyndman, R.J.; Koehler, A.B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679–688.
DOI: <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- [19]. Flores, R.; Valderrama, L. (2021). Evaluation of metallurgical process indicators using percentage-based error metrics. Minerals, 11(8), 864.
DOI: <https://doi.org/10.3390/min11080864>
- [20]. Bendaña, D.; Combes, F. (2016). Performance assessment in flotation circuits using error metrics. Minerals Engineering, 95, 57–66.
DOI: <https://doi.org/10.1016/j.mineng.2016.05.018>
- [21]. ASTM C39/C39M. (2021). Standard Test Method for Compressive Strength of Cylindrical Concrete Specimens. ASTM International.
- [22]. Yeh, I.C. (2006). Analysis of strength of concrete using data mining techniques. Cement and Concrete Research, 36(5), 879–884.
DOI: <https://doi.org/10.1016/j.cemconres.2005.11.005>
- [23]. Hoang, N.D. (2020). Predicting concrete strength using regression and machine learning. Engineering with Computers, 36, 225–235.
DOI: <https://doi.org/10.1007/s00366-018-0660-5>
- [24]. Gandomi, A.H.; Alavi, A.H. (2012). A new multi-gene genetic programming approach to nonlinear system modeling. Engineering Applications of Artificial Intelligence, 25(3), 577–587.
DOI: <https://doi.org/10.1016/j.engappai.2011.09.010>
- [25]. Topçu, I.B.; Saridemir, M. (2008). Prediction of compressive strength of concrete containing fly ash using artificial neural networks. Computers & Materials Science, 41(3), 305–311.
DOI: <https://doi.org/10.1016/j.commatsci.2007.04.009>
- [26]. Erdem, H.; Gül, R. (2012). Prediction of the compressive strength of cement mortar. Construction and Building Materials, 30, 804–812.
DOI: <https://doi.org/10.1016/j.conbuildmat.2011.12.011>
- [27]. Safari, M.; Faradonbeh, R.S. (2018). Reliability-based prediction of concrete strength. Structural Concrete, 19(5), 1425–1436.
DOI: <https://doi.org/10.1002/suco.201700234>
- [28]. Khosravi, K.; Daggupati, P.; Pham, B.T. (2019). Machine learning approaches for predicting concrete strength. Applied Soft Computing, 75, 563–576.
DOI: <https://doi.org/10.1016/j.asoc.2018.11.032>
- [29]. Lin, Y.; Chen, M.; Li, Y. (2020). Data-driven prediction of concrete properties for engineering applications. Automation in Construction, 119, 103325.

DOI: <https://doi.org/10.1016/j.autcon.2020.103325>

[30]. ACI Committee 211. (2002). Standard Practice for Selecting Proportions for Normal, Heavyweight, and Mass Concrete. American Concrete Institute.

DOI: <https://doi.org/10.14359/10987>