

DESARROLLO DE UN MODELO DE MACHINE LEARNING PARA PREDICCIÓN Y AJUSTE DE PARÁMETROS PARA PREVENCIÓN DE ADHERENCIAS EN HORNO ROTATORIO

Development of a Machine Learning Model for Prediction and Parameter Adjustment to Prevent Rotary Kiln Build-up

Daniel Alejandro De la Peña Ibarra^a, Missael Alejandro Lara Castro^a, Jesús Emilio Camporredondo Saucedo^b, Diego Alejandro Tamayo Soriano^c, Alain Rangel Martínez^d, Martín Corona Romo^a, Raul Arellano Aguilar^a.

^a Semillero Internacional de Ingeniería y Procesos Sostenibles (SIIPS), Universidad Politécnica de Monclova Frontera; daniel.delapenaibarra@gmail.com, missael.lara@upmf.edu.mx, martin.corona@upmf.edu.mx, raul.arellano@upmf.edu.mx

^b Facultad de Ingeniería Mecánica y Eléctrica, División de ingeniería mecánica, Universidad Autónoma de Coahuila; emiliocamporredondo@uadec.edu.mx

^c Departamento de Mantenimiento industrial, Universidad Tecnológica de la Región Centro de Coahuila; diego.tamayo@utrcc.edu.mx

^d Corporativo de la Cooperativa la Cruz Azul S.C.L.; alainrangelm@cruzazul.com.mx

Recibido: 12 de febrero 2026. Aceptado: 18 de marzo 2026. Publicado: 23 de abril 2026.

Resumen— La presente investigación se enfoca en desarrollar un modelo predictivo aplicando un modelo de Machine Learning en el proceso de producción de Clínger, específicamente para prever y realizar ajustes de parámetros operativos, reduciendo la formación de anillos y adherencias dentro del horno rotatorio. Para el desarrollo del modelo predictivo se realizó un análisis de correlación de Pearson para identificar los parámetros operativos que contribuyen en mayor grado a la formación de anillos, el cual identificó cinco variables claves divididas en las categorías: química, temperatura interna y alimentación del horno. Una vez determinadas se utilizó un modelo de Árbol de Decisión, y se alimentaron datos operativos. Los resultados demostraron una robusta precisión global del 91% y un sobresaliente recall del 97% para la detección de la formación de un anillo. Este alto recall, garantiza minimizar los Falsos Negativos, considerándose un resultado exitoso y altamente confiable para la alerta temprana, a pesar de la generación de algunas Falsas Alarmas (Falsos Positivos). La integración del modelo en una aplicación interactiva con una capacidad predictiva del 80% en la práctica valida su valor como una herramienta para la toma de decisiones, demostrando que el aprendizaje automático es clave para optimizar la operación de hornos.

Palabras clave— aprendizaje automático, árbol de decisión, formación de anillos y adherencias en horno rotatorio, industria cementera, inteligencia artificial, optimización de procesos.

Abstract— The current investigation focuses on developing a predictive model by applying a machine learning model to the clinker production process, specifically to forecast and adjust operating parameters, reducing the formation of rings and adhesions inside the rotary kiln. To develop the predictive model, a Pearson correlation analysis was performed to identify the operating parameters that contribute most to ring formation, which identified five key variables divided into the categories: chemistry, internal temperature, and kiln feed. Once determined, a decision tree model was used, and operating data was fed into it. The results

demonstrated a robust overall accuracy of 91% and an outstanding recall of 97% for the detection of ring formation. This high recall guarantees the minimization of false negatives, which is considered a successful and highly reliable result for early warning, despite the generation of some false alarms (false positives). The integration of the model into an interactive application with an 80% predictive capacity in practice validates its value as a decision-making tool, demonstrating that machine learning is key to optimizing furnace operation.

Keywords— artificial intelligence, cement industry, decision tree, machine learning, process optimization, ring and adhesion formation in rotary kilns.

I. INTRODUCCION

La industria cementera depende de procesos continuos y altamente controlados, dentro de los cuales el horno rotatorio ocupa un papel fundamental en la producción de clínker. El desempeño eficiente de este equipo resulta crítico, ya que las condiciones de operación determinan la calidad del producto final y la estabilidad de la producción. Uno de los principales problemas que se presentan en este proceso es la formación de adherencias o anillos en el interior del horno (Pisaroni, M., Sadi, R., & Lahaye, D., 2012), lo cual genera variaciones térmicas, obstrucciones en el flujo de material y, en casos severos, paros de producción no planificados.

En este contexto, surge la necesidad de implementar estrategias predictivas que permitan anticipar la formación de adherencias y, además, optimizar los parámetros de operación del horno para reducir su ocurrencia. Una de las herramientas más prometedoras para enfrentar este reto es la inteligencia artificial (IA), particularmente el aprendizaje automático (Machine Learning) (Rojas, E. M., 2020). Mediante el análisis de datos históricos y en tiempo real sobre condiciones de operación, composición de la mezcla cruda, temperaturas y consumo energético, es posible identificar patrones asociados a la formación de anillos y generar predicciones que permitan una intervención temprana y más eficiente.

El presente trabajo de investigación tiene como objetivo desarrollar un modelo de *Machine Learning* capaz de predecir la ocurrencia de adherencias en el horno rotatorio y ajustar parámetros de operación para su prevención. De esta manera, se busca contribuir a la estabilidad del proceso de producción de clínker, reducir los costos derivados de paros no programados y mejorar la eficiencia global del sistema.

II. MATERIALES Y MÉTODOS

A. Software, lenguaje y técnica de correlación utilizados

Para el diseño de este sistema se hizo uso de un lenguaje de programación y los software que se detallan a continuación:

1) Python

Se consideraron diversos lenguajes de programación

ampliamente utilizados en el área industrial y *Machine Learning*, como Java, C++ y Julia, los cuales presentan fortalezas específicas en estadística avanzada, escalabilidad empresarial, alto rendimiento computacional y cómputo científico.

No obstante, se optó por *Python* debido a su ecosistema consolidado de bibliotecas especializadas en aprendizaje automático y redes neuronales, como *TensorFlow*, *PyTorch* y *Scikit-learn*, así como por su sintaxis clara, versatilidad e integración eficiente con lenguajes de bajo nivel. Estas características permiten acelerar los ciclos de desarrollo.

2) Microsoft Excel

Herramienta principal para la gestión y el preprocesamiento de la base de datos del horno rotatorio. Sus funciones clave fueron:

- Almacenamiento y Organización: Recopilación y estructuración de las series temporales de datos operativos.
- Limpieza y Transformación de Datos: Manejo de datos faltantes y la estandarización de variables.
- Generación de Archivos de Entrada: Exportación de las bases de datos limpias y preparadas al formato CSV (Comma Separated Values).

3) Coeficiente de correlación de Pearson

En el análisis estadístico de relaciones entre variables, se consideraron diversos coeficientes de correlación, como el coeficiente de correlación de Spearman y el coeficiente de correlación de Kendall, los cuales ofrecen ventajas en contextos donde existen relaciones monotónicas, datos ordinales o dependencias más complejas. Sin embargo, se seleccionó el *coeficiente de correlación de Pearson* debido a su capacidad para cuantificar de manera precisa la relación lineal entre variables continuas, su alta interpretabilidad y su eficiencia computacional, lo que lo convierte en una herramienta estándar en el análisis exploratorio de datos y en etapas iniciales del desarrollo de modelos de *Machine learning*.

Su ecuación matemática es:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad \text{Ecuación 1}$$

B. Recopilación de datos

Debido a que el objetivo de la presente investigación es encontrar los factores químicos y de proceso con mayor efecto en la formación de adherencias dentro del horno, se llevó a cabo un análisis para la recopilación de estos, destacando 6, los cuales se muestran en la tabla 1. Para cada factor, se desarrollo una base de datos individual, las cuales se alimentaron de las siguientes fuentes de proceso:

- Sensores de temperatura colocados en el exterior del horno rotatorio.
- Sensores de CO y O₂ dentro del horno rotatorio.
- Cálculo de fases y composición química por Bogue.
- Analizador químico por rayos X (FRX) en harina cruda y clínker.
- Reportes de producción por dicho departamento.

Tabla 1.
Contenido de las bases de datos.

Composición química de la harina cruda y Clínker.	Información química: CaO, SiO ₂ , Al ₂ O ₃ , Fe ₂ O ₃ , MgO, SO ₃ , CaF ₂
Composición química y fases calculadas por Bogue	CaO, SiO ₂ , Al ₂ O ₃ , Fe ₂ O ₃ , MgO, C ₃ S, C ₂ S, C ₃ A, C ₄ AF
Temperatura del horno rotatorio y del precalcinador.	Temperaturas de los ciclones anteriores del horno y del mismo.
Concentración de CO y O₂ en el horno rotatorio.	Concentración de CO y O ₂ en el interior del horno rotatorio.
Reportes de producción y combustible.	Tiempo y toneladas de alimentación y producción de Clínker.
Temperatura en la superficie externa del horno rotatorio.	Registro continuo de la temperatura externa del horno rotatorio.
Concentración porcentual de K₂O	Porcentaje contenido de K ₂ O en harina cruda y Clínker.

Las bases de datos contenían 38 variables fisicoquímicas, distribuidas balanceadamente en:

- Químicas: 19 variables
- Físicas: 19 variables

La recolección final de datos constituye un periodo de 6 meses de información (enero 2025 – junio 2025).

C. Selección de variables por criterio cualitativo

Entre las 38 variables recopiladas, se identificó una notable heterogeneidad en los métodos de recolección de datos. A través de un análisis cualitativo, se determinó que el grado de confiabilidad de los datos variaba significativamente, siendo particularmente bajo en aquellos registros provenientes de los sensores de CO y O₂. Esta baja confiabilidad comprometía la validez de los resultados potenciales, justificando la necesidad de un proceso de selección basado en la calidad y la fiabilidad de los datos. Además, en colaboración con los supervisores del área, quienes aportaron su conocimiento empírico, se evaluó la relación potencial entre las variables y la formación de adherencias en el horno rotatorio. Este proceso, permitió identificar las variables con mayor efecto de estar relacionadas con dicho fenómeno.

Al centrarse en las variables que tienen un mayor grado de confiabilidad durante su recolección y una alta probabilidad de estar relacionadas con la formación de adherencias, se optó por utilizar únicamente 22 variables, divididas en las:

- Condiciones operativas, ambientales y la temperatura en los ciclones que preceden al horno.
- Componentes químicos concretos en Clínker y Harina cruda (SO₃ y K₂O)

D. Homologación del formato temporal en las bases de datos

Las variables en las bases de datos tienen frecuencias de registro diferentes o asíncrona como se muestra en la tabla 2.

Tabla 2.
Frecuencia de muestreos de las bases de datos.

Muestreo	Frecuencia (horas)
Composición química de la harina cruda y Clínker.	4
Temperatura del horno rotatorio y del precalcinador.	3
Concentraciones de CO y O ₂ en el horno rotatorio.	2
Reportes de producción y combustible.	8
Temperatura en la superficie externa del horno rotatorio.	8
Concentración porcentual de K ₂ O	24

La frecuencia de registro de datos es asíncrona debido a la organización interna de los departamentos, tiempo que toman los sensores en reflejar los resultados y el tiempo que toman los analizadores químicos generar los resultados.

Se realizó la homologación y selección de las variables más confiables y que mayor efecto tienen (de acuerdo con el personal de supervisión de producción en la empresa) dentro de una sola base de datos con una sola frecuencia de registro: 3 turnos laborados de 8 horas. Y el tipo de valor elegido para las variables fue el promedio y se aplicó de esta forma:

- Las variables que eran registradas por horas, se les calcula el promedio que constituían el turno al que pertenecían.
- Las variables registradas por turnos se aplican sin modificación.
- La variable registrada c/ día se aplica a los 3 turnos que constituyen el formato de 24 horas.

Se realizó el formato de tabla que contendría las 22 variables, colocando el nombre de cada variable como título de columna, la fecha y turno como primeras columnas, formando un valor general para todas las variables y colocando a qué turno pertenece cada variable, siguiendo el horario de cada turno.

Utilizando códigos de programación, se ajustaron los registros correctamente por cada día automáticamente, basándose en un formato de 24 horas por día y como resultado se consigue una base de datos con una sola frecuencia de muestreo: día y turnos, como se muestra en la tabla 3.

Tabla 3.
Fragmento de la base de datos con homogenización temporal.

Fecha	Turno	Tiempo de Op (Hr)	Coque Horno (Toneladas)	Coque Preca (Toneladas)	Promedio de entrada SO ₃ (%)	Promedio Etapa SO ₃ (%)	Promedio en clínker (%)	Promedio K ₂ O
01/01/25	1	8	38.30	40.00	0.28	0.28	1.735	0.7
01/01/25	2	8	38.00	40.00	0.285	0.285	1.88	0.7
01/01/25	3	8	39.00	41.00	0.305	0.305	1.78	0.7
02/01/25	1	8	39.00	39.00	0.33	0.33	1.685	0.72
02/01/25	2	8	38.90	40.00	0.335	0.335	1.86	0.72
02/01/25	3	8	39.50	40.00	0.33	0.33	2.135	0.72
03/01/25	1	8	34.00	36.00	0.335	0.335	1.54	0.8
03/01/25	2	8	34.00	39.00	0.375	0.375	2.08	0.8
03/01/25	3	7.9	35.00	36.00	0.315	0.315	1.975	0.8

E. Identificación y categorización de las filas de datos en las que se registraron formación de adherencias

El nuevo formato de la base de datos ajustada a fechas y turnos contiene 440 filas de datos en los que se detallan las

condiciones fisicoquímicas por cada turno. Para identificar el turno en el que se produjo la formación de una adherencia, se utilizó la base de datos *Registro de temperaturas presentadas en la superficie externa del horno rotatorio*. Esta base de datos proporciona una gráfica de líneas múltiples diariamente que visualiza el perfil de temperatura a lo largo de los 60 metros del horno rotatorio para cada uno de los tres turnos.

En la figura 1 se puede observar el avance de temperatura a lo largo del horno rotatorio durante el primer turno, la cual está dividida en 4 partes principales:

- Entrada de la harina cruda desde el precalcinador. Zona con la menor temperatura dentro del horno; área del metro 44-60.
- “Llantas” o sujeciones exteriores del horno. La temperatura en esta zona es mucho menor debido a que es la parte donde se sujeta el horno y, por lo tanto; es la zona con temperaturas mínimas o valles; área del metro 11 y 43.
- Parte central o media del horno. Zona de mayor temperatura dentro del horno.
- Salida de clínker y fuente de calor alimentando el horno. Zona de alta temperatura dentro del horno; área del metro 1 al 10.

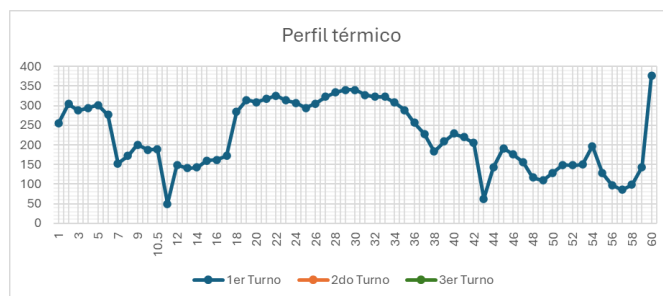


Fig. 1. Perfil térmico del horno rotatorio (1er turno).

La identificación de adherencias en el horno rotatorio a través del perfil térmico externo se fundamenta en la alteración que estas provocan en la transferencia de calor hacia la coraza. En condiciones normales, la temperatura de la coraza en los metros 44 – 60 en la zona de entrada de la harina cruda sigue un patrón relativamente uniforme, determinado por la conductividad del revestimiento refractario y la temperatura interna del horno.

Por otro lado, cuando se desarrollan adherencias en la pared interna, estas forman una capa adicional de material que actúa como aislante térmico, reduciendo el flujo de calor hacia el exterior. Como resultado, en el escaneo infrarrojo de la coraza se observan “valles” o descensos localizados por temperaturas bajas, tal como se muestra en la figura 2 del tercer turno, en los metros 48 y 56.

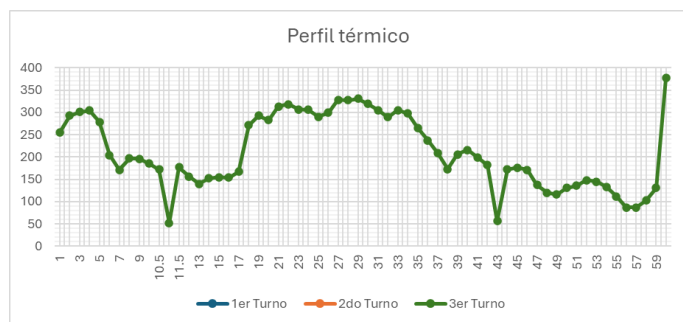


Fig. 2. Perfil térmico con adherencias (3er turno).

La persistencia y localización de estas anomalías térmicas, típicamente en zonas de menor temperatura interna donde los compuestos volátiles tienden a condensarse, permiten justificar científicamente la presencia y el crecimiento de adherencias, diferenciándolas de enfriamientos externos transitorios.

Para categorizar cada uno de los turnos, se identificaron cada día en los que alguno de los turnos generó una adherencia de acuerdo con el perfil térmico y se correlacionó ese día en la base de datos estructurada, permitiendo categorizar cada turno por:

- Formación de adherencias.
- Sin formación de adherencias.

Identificándose:

- 20 días sin formación de adherencias.
- 160 días con formación de adherencias.

F. Aplicación del análisis de correlación de pearson para identificar las variables del proceso y químicas que presentan la mayor correlación con la formación de adherencias

Se aplicó el coeficiente de correlación de Pearson a la base de datos estructurada con la categorización, utilizando un código de programación que lo asigne a los datos considerando sus características propias.

Utilizando la librería *Pandas*, se analiza la correlación entre las variables de proceso y la variable binaria de adherencias. Esto resulta en un valor que indica la fuerza y dirección de la relación: un valor cercano a +1 sugiere que un aumento en el valor de una variable de proceso está fuertemente asociado con la presencia de adherencias, mientras que un valor cercano a -1 sugiere que está asociado con su ausencia. De esta manera, el análisis de Pearson se convierte en una herramienta válida para identificar las variables que ejercen la mayor influencia en la formación de adherencias.

Al ejecutar el código, se exportan los resultados en un archivo csv y se genera un mapa de calor para una mayor facilidad de lectura, como se muestra en el siguiente mapa de calor de la figura 3.

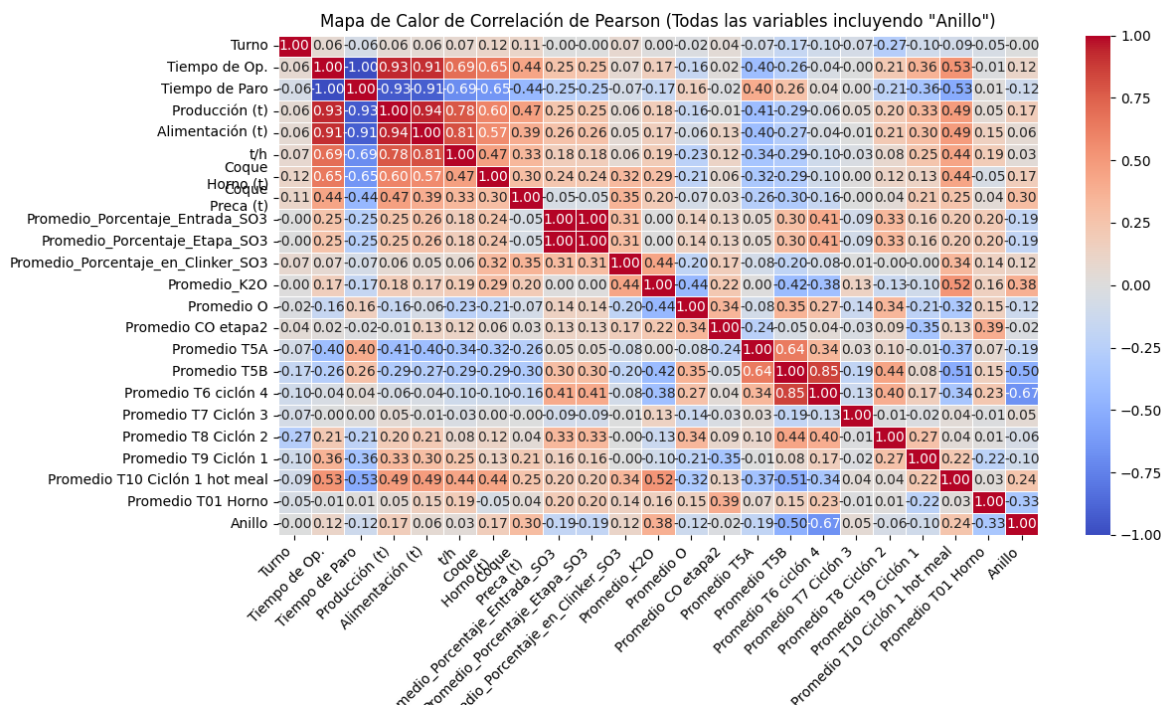


Fig. 3. Mapa de calor de Correlación de Pearson.

El mapa de calor de Pearson se lee a través de su escala de colores, la cual representa la fuerza y dirección de la correlación entre las variables. Cada celda del mapa, ubicada en la intersección de una fila y una columna, muestra el coeficiente de correlación (r) entre esas dos variables. Los colores cálidos (como el rojo) y los valores positivos (cerca de +1) indican una correlación positiva fuerte, lo que significa que a medida que una variable aumenta, la otra también tiende a aumentar. Por el contrario, los colores fríos (como el azul) y los valores negativos (cerca de -1) señalan una correlación negativa fuerte, donde el aumento de una variable se asocia con la disminución de la otra. Finalmente, los colores neutros (blanco o gris) y los valores cercanos a 0 sugieren una correlación lineal débil o nula.

Al seleccionar únicamente las 5 variables con la mayor correlación con la formación de adherencias, se genera una nueva base de datos, como se muestra en la tabla 4.

Tabla 4.
Variables con la mayor correlación con la formación de adherencias.

Fecha	Coque Horno	Coque Preca (Toneladas)	Promedio K ₂ O	Promedio T5B (°C)	Promedio T6 Ciclón 4 (°C)	Anillo (1= Sí, 0 = No)
15/02/25	43	31	0.96	397	551.5	1
18/02/25	40	34	0.85	408	564	1
01/03/25	40	34	1.09	420.5	583.5	0
07/01/25	42	35	0.74	422	583	1
18/02/25	42	35	0.85	404.5	569.5	1
28/02/25	42	35	1.1	426.33	591.33	1
28/02/25	42	35	1.1	426.33	591.33	0
24/02/25	41	35	0.91	403.5	567.5	0

Esta base de datos contiene un total de 219 filas de datos, incluyendo turnos con formación y sin formación de anillos, los

cuales son identificados por la fecha que le concierne a cada uno de ellos.

G. Selección del tipo de modelo de machine learning

La elección del modelo de Machine Learning se basó en una evaluación de las características de la base de datos y los requisitos del problema de predicción. Dado que la tarea es de clasificación binaria (formación de adherencia o no formación), se consideraron diversos algoritmos como regresión logística, máquinas de vectores de soporte (SVM) y árboles de decisión.

El árbol de decisión se seleccionó como el modelo principal (Asanza, W. R., & Olivo, B. M., 2018) por tres razones clave.

- Su naturaleza intuitiva y visual facilita la interpretación del modelo, permitiendo comprender de manera clara qué variables y en qué umbrales influyen en la predicción.
- Los árboles de decisión no requieren una normalización rigurosa de los datos, lo que agiliza el proceso de pre-procesamiento.
- Son particularmente efectivos para capturar relaciones no lineales y complejas entre las variables, una característica relevante en el comportamiento impredecible de las adherencias en los hornos.

Estas razones permiten no solo predecir la ocurrencia de adherencias, sino también identificar las reglas y puntos críticos del proceso que las originan (UptimeAI, 2022), tal como se muestra en el árbol desarrollado en la figura 4, en donde siguiendo las condiciones de decisión se puede determinar a qué clasificación pertenece (formación o no formación de adherencias y anillos).

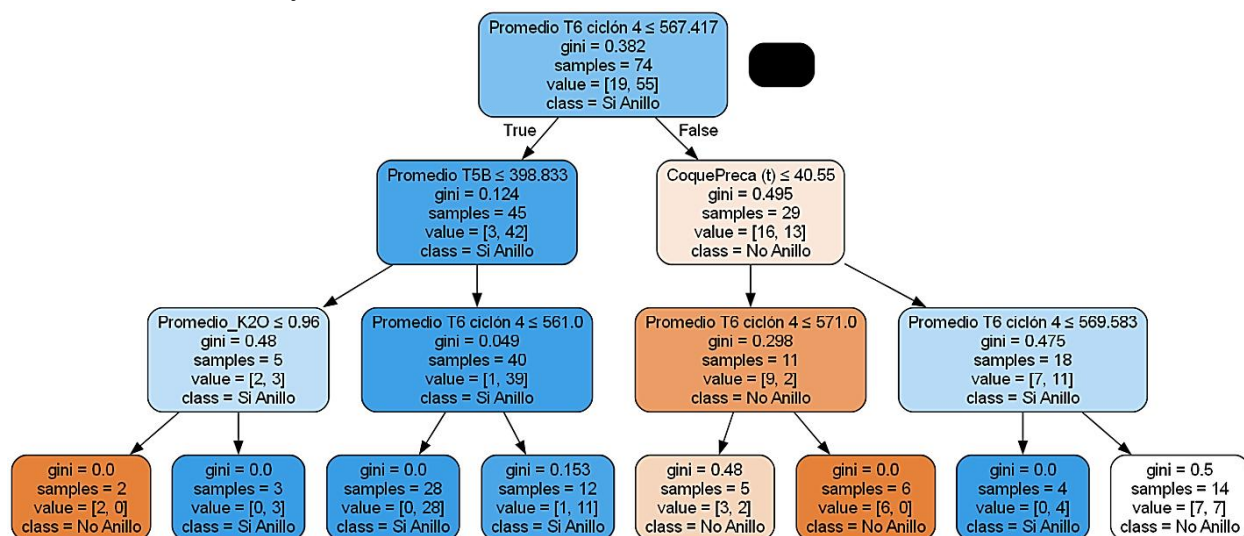


Fig. 4. Ejemplo del Modelo de Árbol de decisión.

H. Desarrollo y entrenamiento

El árbol de decisión utilizado para este proyecto fue programado utilizando las librerías de Python *Scikit-learn* y *Matplotlib*, lo que permitió no solo entrenar el modelo con los datos históricos del proceso, sino también visualizar y validar las reglas de clasificación aprendidas por el algoritmo. Esta fase de programación fue crucial para transformar las correlaciones identificadas en un sistema predictivo funcional y transparente.

Su objetivo principal es generar un modelo predictivo capaz de identificar la formación de adherencias (Anillo) a partir de variables operativas del proceso (FEATURES), como se explica a continuación:

1) Carga y preparación de datos

El primer paso es importar las librerías necesarias, como *Pandas* para la manipulación de datos y *Scikit-learn* para el desarrollo del modelo. El script lee un archivo CSV y define las variables predictoras (FEATURES) y la variable objetivo (TARGET_COLUMN). Es crucial que la variable objetivo, Anillo, se convierta a un formato numérico (0 y 1) para que el algoritmo pueda procesarla, como se muestra en la figura 5.

Se evaluó el tratamiento de datos faltantes mediante el método *dropna()*. Aunque existen técnicas de imputación (como la media o interpolación lineal), se optó por la eliminación de registros incompletos para preservar la integridad física de las lecturas. En procesos industriales sensibles, la imputación puede introducir ruido o sesgos artificiales que oculten las correlaciones reales entre las temperaturas extremas y la formación de anillos. Tras la limpieza, el dataset conservó 174 muestras representativas, volumen suficiente para garantizar la convergencia del algoritmo de árbol de decisión sin comprometer la precisión con datos estimados.

```
16 # --- Cargar y Preparar los Datos ---
17 data = pd.read_csv(CSV_FILE_PATH)
18 X = data[FEATURES]
19 y = data[TARGET_COLUMN]
20
21 if not pd.api.types.is_numeric_dtype(data[TARGET_COLUMN]):
22     data[TARGET_COLUMN] = pd.to_numeric(data[TARGET_COLUMN])
23     data.dropna(subset=[TARGET_COLUMN], inplace=True)
```

Fig 5. Carga y preparación de datos.

2) División de datos

Para asegurar que el modelo sea evaluado de manera imparcial y no haya memorizado los datos de entrenamiento, el conjunto de datos se divide en dos partes:

- Conjunto de entrenamiento (X_{train} , Y_{train}): Contiene el 80% de los datos, utilizados para que el modelo

"aprenda" las relaciones entre las variables y la formación de adherencias.

- Conjunto de prueba (X_{test} , Y_{test}): Contiene el 20% restante de los datos. Estos datos son "invisibles" para el modelo durante el entrenamiento y se utilizan para evaluar su rendimiento y precisión.

En la figura 6 se muestra la sección en donde se diferencian los datos de entrenamiento y prueba.

```
1 # --- Dividir Datos en Entrenamiento y Prueba ---
2 X_train, X_test, y_train, y_test = train_test_split(
3     X, y, test_size=0.20, random_state=42, stratify=y
4 )
```

Fig. 6. Estipulación de los datos de entrenamiento y prueba.

3) Entrenamiento del modelo

El núcleo del proceso es el entrenamiento del árbol de decisión. El código crea una instancia del clasificador (*DecisionTreeClassifier*) y especifica sus parámetros clave, en la figura 7 encontramos las especificaciones técnicas en el código:

- *max_depth*: Limita la profundidad del árbol para prevenir el sobreajuste.
- *criterion='entropy'*: Define el método para medir la pureza de los nodos, buscando maximizar la información ganada en cada división.
- *min_impurity_decrease* y *ccp_alpha*: Son parámetros de poda que ayudan a simplificar el árbol, evitando que se vuelva demasiado complejo y mejoren su generalización. Finalmente, el método *fit()* entrena el modelo usando los datos de entrenamiento.
- *ccp_alpha=0.01*: Este parámetro aplica una poda de complejidad de costos que penaliza la creación de nodos innecesarios. Su función es simplificar la estructura del árbol para evitar el sobreajuste, asegurando que el modelo identifique patrones operativos reales en lugar de memorizar datos específicos del periodo histórico.

```
1 # --- Entrenar el Modelo de Árbol de Decisión ---
2 model = DecisionTreeClassifier(
3     max_depth=MAX_TREE_DEPTH,
4     random_state=42,
5     criterion='entropy',
6     min_impurity_decrease=0.01,
7     ccp_alpha=0.01
8 )
9 model.fit(X_train, y_train)
```

Fig. 7. Entrenador del modelo de Machine Learning.

4) Guardado y exportación del modelo

Una vez entrenado, el modelo es guardado en formato joblib para su posterior uso. En la figura 8 se presenta la conformación de las instrucciones de almacenamiento.

```
1 # --- Guardar el modelo en un archivo .joblib ---
2 output_model_file = 'Arbol_desicion.joblib'
3 joblib.dump(model, output_model_file)
```

Fig. 8. Guardar el modelo de IA.

5) Evaluación del rendimiento

Para validar la efectividad del modelo, se evalúa su rendimiento en el conjunto de prueba. El código genera las siguientes métricas visuales en la terminal de VSC:

- accuracy_score
- classification_report
- confusion_matrix

6) Visualización del árbol

Como parte final del código, se utilizó la librería *Pydotplus*, convierte el modelo en una imagen visual que representa el árbol de decisión con nodos y ramas.

I. Diseño visual de la aplicación

Antes de la programación de la aplicación para integrar el modelo predictivo, se realizó el diseño visual que se desea que tenga la aplicación, el resultado esperado se realizó con figuras en Microsoft PowerPoint, tal como se muestra en la figura 9.

Sistema de Predicción de adherencias

Ingresar los Parámetros del Proceso

Parámetro:	Valor:
Parámetro:	Valor:
Parámetro:	Valor:
Parámetro:	Valor:

Predicción y Recomendación:

Visualizar árbol de decisión

Resultados de la Predicción

Predicción:

Área de sugerencias

Fig. 9. Diseño visual de la aplicación.

J. Programación e implementación del modelo de machine learning en la aplicación interactiva

Para programar la aplicación se designan las 4 funciones lógicas que tiene que poseer:

- Predecir adherencias
- Sugerir ajustes en parámetros
- Rangos de datos
- Regla de proceso

K. Método de validación aplicado

Para comprobar la efectividad y valorar el rendimiento de la aplicación, se llevó a cabo una serie de pruebas utilizando registros pasados del proceso de clinkerización.

Se compiló un conjunto de datos a partir de múltiples registros pasados. Cada uno de estos registros correspondía a un día y turno específico en el historial del proceso de clinkerización, detallando los parámetros operacionales de ese día y si hubo o no generación de adherencias. Posteriormente, la aplicación fue utilizada para predecir el estado de cada uno de estos días históricos. El proceso de validación de la aplicación se efectuó mediante el registro sistemático en la tabla de la exactitud de cada predicción, contrastando el resultado arrojado por la aplicación con el evento real documentado.

III. RESULTADOS

1) Criterio de formación de adherencias a través del análisis del perfil térmico del horno rotatorio

En la zona de entrada del horno, donde las temperaturas oscilan entre 800 y 1100 °C, estos elementos se combinan para formar sales; como sulfatos y cloruros alcalinos (K_2SO_4 , Na_2SO_4 , KCl , $NaCl$), las cuales poseen puntos de fusión relativamente bajos, además, presentan fuerte tendencia a la volatilización/condensación. Cuando estos vapores migran hacia zonas más frías del horno, se condensan sobre la harina cruda o directamente sobre la superficie del refractario, generando una matriz pegajosa que actúa como aglutinante de partículas, además, la plasticidad transitoria de minerales intermedios como la belita (C_2S) y los aluminatos, en rango térmico medio (el intervalo de temperatura en el que los minerales adquieren una consistencia pegajosa, generalmente entre los 600 y 900 °C) adquieren propiedades semisólidas incrementando las condiciones para la formación de anillos y adherencias.

La fundamentación química y térmica son criterios esenciales

para determinar la formación de anillos y adherencias dentro de un horno cementero. Por lo que en base al análisis del perfil térmico de este se plantean los siguientes criterios que promueven dicho fenómeno: se consideró que existe formación de anillos y adherencias cuando la temperatura exterior de la coraza es menor o igual a 150 °C, entre los metros 44 al 59 del horno.

2) Análisis de los coeficientes de correlación de pearson

Del mapa se pueden extraer las 5 variables con una mayor correlación con la formación de adherencias (Mun, S., 2024) con un coeficiente c/u de:

- Coque en el horno (T): 0.17
- Coque en el precalcinador (T): 0.30
- Potasio (K_2O) (%): 0.38
- Temperatura en T5B (°C): -0.50
- Temperatura en T6 Ciclón 4 (°C): -0.67

Significado de cada coeficiente:

- Potasio (K_2O) (%): $r=0.38$: Valor positivo más alto, sugiere una correlación moderada y positiva con la formación de adherencias. El potasio es un álcali que se volatiliza en las zonas calientes del horno y se condensa en las zonas más frías, formando una fase líquida y pegajosa que actúa como un aglutinante para las partículas, iniciando la formación de adherencias (McGinley, M. W. *et al.*, 2024).
- Coque en el precalcinador (T): $r=0.30$: Valor que indica una correlación débil a moderada y positiva. Un aumento en la cantidad de coque en el precalcinador se asocia con un aumento en la probabilidad de formación de adherencias.
- Coque en el horno (T): $r=0.17$: La correlación con el coque en el horno es muy débil. Su bajo valor sugiere que esta variable tiene una influencia limitada en la formación de adherencias en comparación con otras.
- Temperatura en T5B (°C): $r=-0.50$: Coeficiente negativo moderadamente fuerte. Indica una relación inversa: a medida que la temperatura en este punto del horno (T5B) disminuye, la probabilidad de formación de

adherencias aumenta.

- Temperatura en T6 Ciclón 4 (°C): $r=-0.67$: Correlación fuerte. Indica una relación negativa y fuerte. Una caída en la temperatura de este ciclón (T6) está fuertemente relacionada con un aumento en la probabilidad de formación de adherencias.

3) Evaluación del árbol de decisión obtenida durante el entrenamiento

Precisión Global (Accuracy): 0.91. El modelo clasificó correctamente el 91% de todas las instancias en el conjunto de prueba. La tabla 5 muestra los resultados desglosados:

Tabla 5.
Resultados de la evaluación del entrenamiento del Árbol.

Clase	Precisión	Sensibilidad (Recall)	Puntuación F1 (F1-score)	Soporte
No Anillo	0.90	0.75	0.82	12
Si Anillo	0.91	0.97	0.94	32
Promedio	0.91	0.91	0.91	44

Se implementó la métrica *class_weight='balanced'* para equilibrar las dos clases y que el entrenamiento no se incline hacia la clase mayoritaria, proporcionando un entrenamiento con mayor grado de precisión entre las dos partes de manera equitativa

4) Árbol de decisión generado

El árbol generado se observa en la figura 10 con las métricas de evaluación más altas, a su vez, es el selecto para utilizarse para predecir adherencias.

Para dar soporte matemático al modelo, se cuantificó la importancia de las características basándose en la entropía. Se determinó que la variable Promedio T6 ciclón 4 ostenta el 41% del peso predictivo, como se visualiza en la figura 11, consolidándose como el indicador crítico para la detección de anillos, por encima de los niveles de potasio o coque.

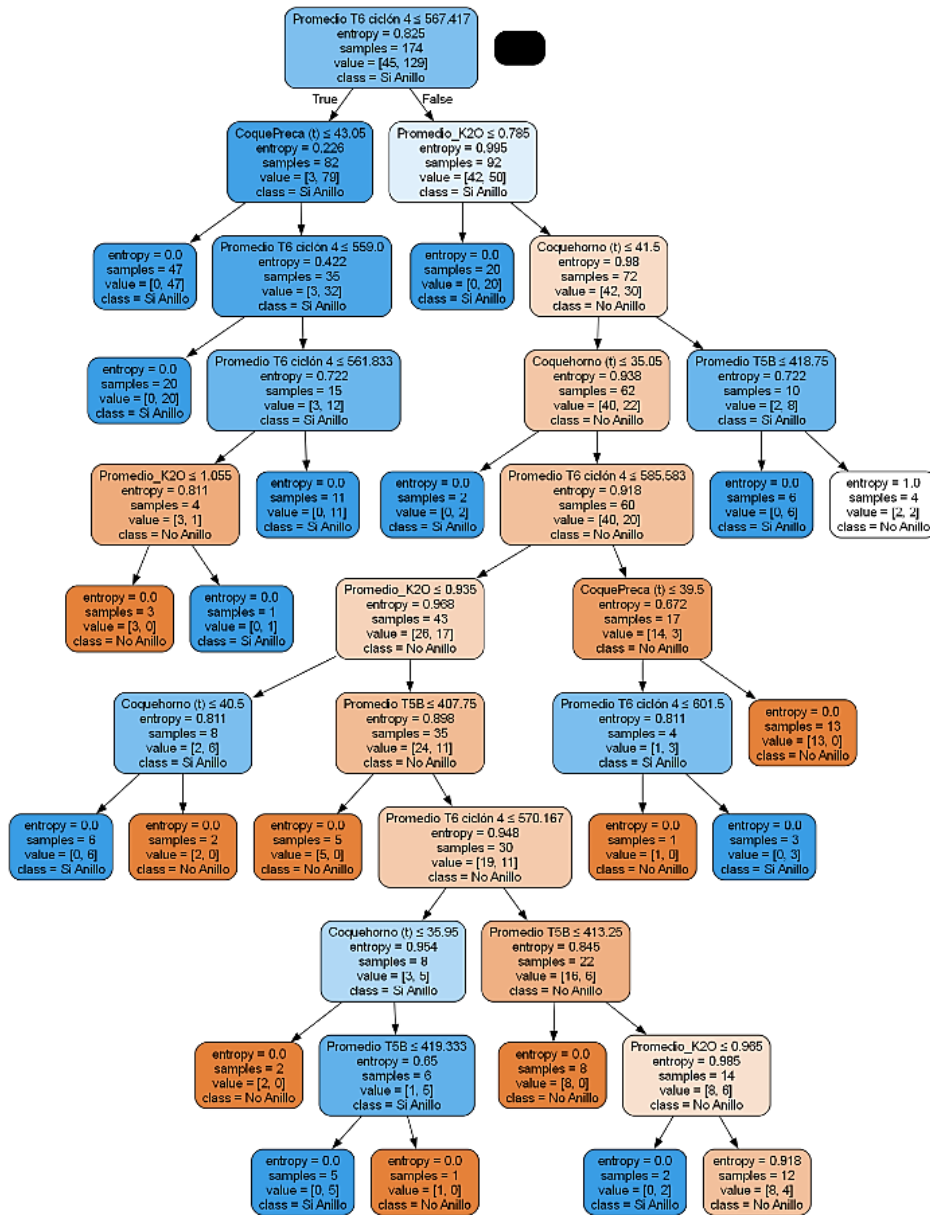


Fig. 10. Árbol de decisión usado para el tema de investigación.

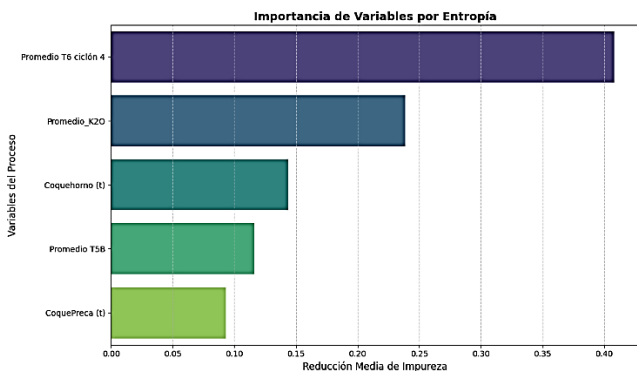


Fig. 111. Análisis cuantitativo: Importancia de variables por Entropía.

5) Análisis de la aplicación interactiva integrando el modelo predictivo

Considerando la lógica interna de las funciones principales, sumando los elementos visuales para la interacción del usuario con la lógica, se desarrolla la aplicación interactiva presentada en la figura 12.

Fig. 122. Aplicación: Sistema de Predicción de Adherencias en Horno Rotatorio.

Al utilizarse la aplicación se notan 3 tipos de respuestas:

- Formación de adherencias (Con sugerencias de ajuste para evitar la formación): Cuando la combinación de variables resulta en una formación de adherencias y anillos la aplicación nos presenta las sugerencias de ajuste que pueden realizarse para evitar la formación.
- No se detecta formación de adherencias: Cuando la aplicación no predice la formación del fenómeno de acuerdo con las variables ingresadas.
- Formación de adherencias (Sin sugerencias de ajuste para evitar la formación): La aplicación predice la formación de adherencias y anillos sin encontrar ajustes en los parámetros para prevenirla.

6) Validación y evaluación del desempeño de la aplicación al realizar predicciones y sugerir ajustes

Al evaluar y validar el desempeño de la aplicación al predecir adherencias, el reporte 'Evaluación de la capacidad predictiva y rendimiento de la aplicación' obtuvo los siguientes resultados:

- Predijo correctamente la clasificación de 16/20 registros de adherencias, obteniendo una evaluación de la capacidad predictiva del 80% en este grupo de 20 días.
- Registró 4 falsos negativos como única predicción equivocada.
- Al analizar las fechas que la aplicación predijo erróneamente en el reporte, se encontró que el día 27/04/2025, turno A y C, tenía un pronunciado valle en

el metro 58 a 60 aproximadamente. Aunque no tenía las condiciones necesarias para generar adherencias en ese momento, falló la predicción al ser este generado con antelación desde el día 9 de abril y continuando hasta prevalecer en las fechas 22/05/2025 y 25/05/2025.

- La aplicación no muestra sugerencias de ajustes cuando los parámetros registrados están muy inclinados a formar adherencias (Benchekroun, M. T., Zaki, S., Hezzem, B., & Laacha, H., 2023).

IV. CONCLUSIONES

A. Correlación de variables clave en la formación de adherencias.

- El análisis de correlación de Pearson identificó las variables con mayor relación con la formación de adherencias en el horno rotatorio, alineándose con la fundamentación teórica del proceso.
- Variables con Correlación Fuerte y Moderada: La temperatura en T6 Ciclón 4 (-0.67) y la temperatura en T5B (-0.50) muestran una fuerte y moderada correlación negativa, respectivamente. Esto confirma que una disminución en las temperaturas en las zonas de precalentamiento y ciclones está directamente relacionada con la formación de adherencias, debido a la condensación de compuestos volátiles.
- Variables con Correlación Positiva: La concentración de Potasio (0.38) en la materia prima tiene una correlación moderada y positiva. Este resultado valida la teoría de que los compuestos alcalinos son precursores clave de las adherencias. Asimismo, el Coque en el precalcinador (0.30) presenta una correlación débil a moderada, mientras que el Coque en el horno (0.17) muestra una correlación muy débil. Estos resultados sugieren que el control de la temperatura y la concentración de potasio son los factores más influyentes para predecir la formación de adherencias.

B. Desempeño del modelo de árbol de decisión

El modelo de Árbol de Decisión ha demostrado ser una herramienta altamente eficaz y confiable para predecir la formación de adherencias:

- Precisión del Modelo: Con una precisión global del 91% en el conjunto de prueba, el modelo muestra una robusta capacidad para clasificar correctamente los casos.
- Detección de Anillos (Recall): El Recall de 0.97 para la clase "Si Anillo" representa el hallazgo más significativo de la investigación. Este valor indica que

el modelo identifica correctamente la formación de adherencias en el 97% de los casos, fallando en solo uno de los 32 eventos de prueba.

Esta alta sensibilidad se priorizó frente a la precisión absoluta debido a la asimetría de costos en la operación del horno. Mientras que un Falso Positivo (falsa alarma) solo conlleva ajustes operativos menores sin interrumpir la marcha, un Falso Negativo (no detectar un anillo) puede derivar en una obstrucción crítica y un paro no programado. El costo de detener y recalentar el horno, sumado a la pérdida de producción y posibles daños al refractario, justifica un enfoque preventivo donde el modelo actúe como un sistema de alerta temprana de confiabilidad, priorizando la continuidad del proceso sobre la evitación de ajustes innecesarios.

C. Aplicación interactiva y validación

- La implementación de una aplicación interactiva que integra el modelo predictivo valida su utilidad en un entorno operativo real.
- Capacidad Predictiva en Campo: La aplicación demostró una capacidad predictiva del 80% en un conjunto de 20 días evaluados (Chatti, S., & Ghorbel, A., 2021).
- Errores Específicos: El análisis de los cuatro falsos negativos, iniciando con el día 27/04/2025 (donde existía un anillo, pero la aplicación no lo detectó) reveló la detección de un solo valle en los metros 58 a 60 del horno formado con semanas de antelación, fundamentando que, aunque los parámetros no favorecen la formación de adherencias, un fenómeno formado con mucho tiempo se vuelve difícil de deshacer por condiciones fisicoquímicas internas del horno. Aun cuando este no tiene las condiciones para formarse, no se deshace.
- Funcionalidad y usabilidad: Su incapacidad para proporcionar sugerencias en casos extremos (cuando los parámetros están fuertemente inclinados a formar adherencias) destaca una oportunidad de mejora y una alarma para el departamento de producción para detectar combinaciones de parámetros que tengan una alta probabilidad de formar adherencias y evitarlas.

En conclusión, el modelo predictivo y su aplicación interactiva representan una herramienta valiosa para la gestión proactiva del proceso de producción en el horno rotatorio. Ha demostrado ser robusto y altamente efectivo para detectar el fenómeno de adherencias, lo que contribuye a la optimización de la operación, la reducción de interrupciones en la producción no programados y el aumento de la eficiencia del proceso.

1. Agradecimientos

Los autores agradecen a la empresa Cementos Cruz Azul Aguascalientes por permitir y facilitar la información; además por la estadía en sus instalaciones para la realización de esta investigación.

Finalmente se agradece a la señora Laura Ibarra y a la Ingeniera Arianna Ibarra por su apoyo económico para la solvencia del necesario viaje nacional para llevar a cabo este proyecto.

2. Referencias

- [1] Rojas, E. M. (2020). Machine Learning: análisis de lenguajes de programación y herramientas para desarrollo. *Revista Ibérica de Sistemas e Tecnologías de Informação*, (E28), 586-599. <https://www.proquest.com/openview/c7e24c997199215a26a39107dd2fe98/1?pq-origsite=gscholar&cbl=1006393>
- [2] Sarmiento-Ramos, J. L. (2020). Aplicaciones de las redes neuronales y el deep learning a la ingeniería biomédica. *Revista UIS Ingenierías*, 19(4), 1-18. <https://www.redalyc.org/journal/5537/553768213002/553768213002.pdf>
- [3] Asanza, W. R., & Olivo, B. M. (2018). Redes neuronales artificiales aplicadas al reconocimiento de patrones. *Editorial UTMACH*, 1(4), 5. https://www.researchgate.net/profile/Bertha-Mazon-Olivo/publication/327703478_Capitulo_1_Generalidades_de_las_redes_neuronales_artificiales/links/5b9fe3c0299bf13e6038a1d8/Capitulo-1-Generalidades-de-las-redes-neuronales-artificiales.pdf?__cf_chl=tk=XWjVkp7ffnnKTENm.ySLHz6Zzxnc1jbfT.6HAh22RM-1757631095-1.0.1.1-wLryd7fuTua.AO_oVO50MP1gLf9CGilasObj9DLP_78
- [4] Pisaroni, M., Sadi, R., & Lahaye, D. (2012). Counteracting ring formation in rotary kilns. *Journal of Mathematics in Industry*, 2(1), Article 3. <https://doi.org/10.1186/2190-5983-2-3>
- [5] Mun, S. (2024). Operating Key Factor Analysis of a Rotary Kiln Using a SHAP-based Explainable Artificial Intelligence Approach. *Electronics*, 13(22), 4413. <https://www.mdpi.com/2079-9292/13/22/4413>
- [6] Benchekroun, M. T., Zaki, S., Hezzem, B., & Laacha, H. (2023). Kiln Process Fan Vibrations Prediction Based on Machine Learning Models: Application to the Raw Mill Fan. *Computer Sciences & Mathematics Forum*, 6(1), 6. <https://doi.org/10.3390/cmsf2023006006>
- [7] Ali, A. M., Tabares, J. D., & McGinley, M. W. (2022). A machine learning approach for clinker quality prediction and nonlinear model predictive control design for a rotary cement kiln. *Journal of Advanced Manufacturing and*

- Processing, 4(4), e10137.
<https://doi.org/10.1002/amp2.10137>
- [8] McGinley, M. W., et al. (2024). Industrial-scale Prediction of Cement Clinker Phases using Machine Learning and NMPC design. arXiv. <https://arxiv.org/abs/2412.11981>
 - [9] Machalek, D. (2019). Model predictive control of a rotary kiln for fast electric.... Computers & Chemical Engineering. <https://www.sciencedirect.com/science/article/am/pii/S0892687519304327>
 - [10] UptimeAI. (2022). Transforming Rotary Kiln Maintenance Workflows With Predictive Analytics. UptimeAI. <https://www.uptimeai.com/resources/kiln-predictive-maintenance>
 - [11] Shovlin, J., Liu, K., Shen, Y., Borghard, B., Makse, H. A., & Tomassone, M. S. (2022). Discrete Element Simulations and Machine Learning for Improving the Performance of Dry Catalyst Continuous Impregnation Processes. arXiv preprint arXiv:2212.00634. <https://arxiv.org/abs/2212.00634>
 - [12] Svensen, J. L., da Silva, W. R. L., Pigazo Merino, J., Sampath, D., & Jørgensen, J. B. (2024). A Dynamical Simulation Model of a Cement Clinker Rotary Kiln. arXiv preprint arXiv:2405.03200. <https://arxiv.org/abs/2405.03200>
 - [13] Chatti, S., & Ghorbel, A. (2021). Analysis and prediction of ring formation in a rotary cement kiln using artificial neural networks. Journal of Applied Systems Innovation, 4(2), 28. <https://doi.org/10.3390/asi4020028>
 - [14] Li, X., Zhang, H., & Liu, Q. (2022). Real-time monitoring and prediction of ring formation in cement kilns based on computer vision and thermal imaging. Case Studies in Thermal Engineering, 35, 102115. <https://doi.org/10.1016/j.csite.2022.102115>
 - [15] Gao, J., & Wang, Y. (2023). Optimization of Cement Kiln Parameters Using Deep Reinforcement Learning. Journal of Process Control, 124, 45-58. <https://doi.org/10.1016/j.jprocont.2023.02.004>