

Design of a SCADA system for predictive maintenance using machine learning in a TR injection machine

Guzman Brayan¹; Arias Jeanpierre²; Alfaro Anthony³

^{1,2,3}Universidad Tecnológica del Perú, Perú, U19207487@utp.edu.pe, 1320892@utp.edu.pe, c22020@utp.edu.pe

Abstract– *This study proposes the design of a SCADA system incorporating Machine Learning and linear regression to enhance predictive maintenance for the TR injection molding machine. The objective is to minimize failures and operational costs through the analysis of electrical and environmental parameters that influence the machine's performance. The system's effectiveness is evaluated via simulations, with an emphasis on failure prediction and maintenance optimization.*

Keywords– *SCADA, machine learning, TR injection molding machine, predictive maintenance*

Diseño de un sistema SCADA para mantenimiento predictivo utilizando machine learning en una máquina inyectora TR.

Guzman Brayan¹; Arias Jeanpierre²; Alfaro Anthony³

^{1,2,3}Universidad Tecnológica del Perú, Perú, U19207487@utp.edu.pe, 1320892@utp.edu.pe, c22020@utp.edu.pe

Resumen– Este estudio propone diseñar un sistema SCADA con Machine Learning y regresión lineal para mejorar el mantenimiento predictivo de la máquina inyectora TR. Se busca reducir fallos y costos operativos mediante el análisis de parámetros eléctricos y ambientales que afectan su rendimiento. La efectividad del sistema se evalúa a través de simulaciones, enfocándose en la anticipación de fallos y optimización del mantenimiento.

Palabras clave– Scada, machine learning, inyectora TR, mantenimiento predictivo

I. INTRODUCCIÓN

La eficiencia y la confiabilidad de las máquinas industriales son fundamentales para optimizar la productividad y la rentabilidad en la fabricación de calzado. En este sentido, las máquinas inyectoras de plástico juegan un papel crucial, ya que permiten la producción de suelas de calzado en material TR (goma termoplástica). Estas máquinas se componen de tres partes esenciales: la unidad de inyección, la unidad de cierre y el sistema de control. En algunos casos se incluye un sistema de resistencias eléctricas para la inyección de material fundido que permiten calentar las boquillas a una determinada temperatura cuyo funcionamiento adecuado es vital para alcanzar las temperaturas óptimas de operación [1],[2].

La falta de un sistema efectivo de monitoreo para las resistencias eléctricas en la máquina inyectora TR se traduce en un incremento de la frecuencia de fallos y la necesidad de intervenciones correctivas. Este escenario no solo eleva los costos operativos, sino que también compromete la continuidad de la producción, afectando la competitividad de la planta. De este modo, las empresas que adoptan el mantenimiento predictivo pueden evitar accidentes, reducir drásticamente el tiempo de inactividad de los equipos y mejorar los procesos, lo que pone de manifiesto la relevancia de contar con herramientas modernas que faciliten la gestión proactiva del mantenimiento [3].

La investigación sobre el mantenimiento predictivo ha demostrado que los sistemas avanzados de monitoreo, como el desarrollo de sistemas SCADA (Supervisory Control and Data Acquisition) se vuelve imperativo. Estos sistemas no solo permiten la visualización y registro en tiempo real del comportamiento de las resistencias eléctricas, sino que también pueden alimentarse de datos a través de instrumentos de medición que utilizan técnicas de Machine Learning para optimizar la toma de decisiones [4].

La presente investigación tiene como objetivo principal determinar cómo el diseño de un sistema SCADA utilizando regresión lineal y Machine Learning como técnica clave de análisis, puede revolucionar la gestión del mantenimiento predictivo en una máquina inyectora TR. Este enfoque innovador aprovecha la regresión lineal para identificar y modelar relaciones entre variables operativas, lo que permite predecir con precisión posibles fallos y patrones de desgaste. De este modo, no solo se busca optimizar la operatividad de la máquina, sino también minimizar los costos operativos asociados a fallos y mantenimientos correctivos, garantizando una gestión más eficiente y basada en datos. El Machine Learning, se emplea para predecir un resultado, que es necesario analizar, para observar si los valores que produce durante el entrenamiento se comparan con algunos valores objetivos [5]. Además, se propone realizar un análisis exhaustivo de los parámetros eléctricos y ambientales que influyen en el rendimiento de la máquina. Esto permitirá identificar los factores críticos que afectan su funcionamiento y, a su vez, facilitará el diseño de un modelo de gestión de mantenimiento predictivo adaptado a las necesidades específicas de la empresa. Para poder completar esta información también es necesario recopilar datos de máquina, tales como datos detallados del proceso, situación de condición de ciertos componentes críticos [6].

Finalmente, se evaluará la efectividad del sistema propuesto con simulaciones en el entorno SCADA y Machine Learning, validándolo mediante métricas específicas y validación cruzada; analizando cómo estas herramientas pueden mejorar la capacidad de anticipación ante fallos y optimizar los procesos de mantenimiento. El artículo está organizado de la siguiente manera: en el Capítulo I se presenta la introducción; el Capítulo II aborda la metodología empleada en el proceso SCADA y Machine Learning; en el Capítulo III se detalla el diseño del sistema SCADA, así como la supervisión y adquisición de datos. De la misma manera en este capítulo se muestran el HMI. El Capítulo IV se centra en la elección del modelo de Machine Learning, la validación y el entrenamiento de este modelo. En el Capítulo V se presentan los resultados obtenidos, el Capítulo VI ofrece las conclusiones del estudio y finalmente en el capítulo VII se tiene las recomendaciones que hacemos para futuras investigaciones o mejoras sobre el trabajo realizado.

II. METODOLOGÍA

Para garantizar la eficacia del proyecto, se optó por emplear la metodología ágil ya que se caracteriza por su enfoque iterativo, flexible y centrado en el usuario, lo que la convierte en una excelente opción para proyectos que requieren adaptación constante y entregas incrementales. Este enfoque permite trabajar en ciclos cortos, llamados sprints, lo que facilita la incorporación de cambios y la mejora continua del producto. En el caso del diseño del sistema SCADA, donde su principal función es generar una base de datos para almacenar información, así como mostrar en una interfaz las fechas de mantenimiento y mantener un histórico de eventos, la metodología ágil fue ideal. A través de iteraciones, se pudo desarrollar y probar la base de datos y la pantalla visual, ajustando cada componente según la retroalimentación de los usuarios. Esto permitió asegurar que el sistema se adaptara de manera eficiente a las necesidades operativas reales, optimizando recursos, reduciendo tiempos de desarrollo y garantizando un diseño funcional y accesible que cumpliera con los objetivos de mantenimiento predictivo.

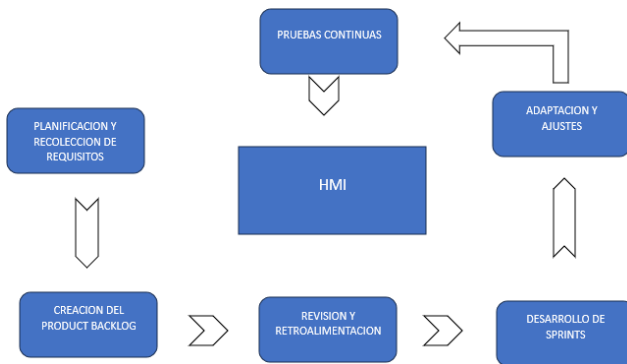


Fig. 1 Diagrama del proceso SCADA.

En el desarrollo de Machine Learning, la metodología ágil facilita un enfoque iterativo, permitiendo la construcción incremental del modelo y la optimización continua mediante ajustes rápidos basados en métricas de desempeño y retroalimentación constante. Para este proyecto de predicción de fallas, resulta fundamental la adquisición de datos confiables, obtenidos a través de sensores seleccionados de manera estratégica en función de su precisión y relevancia para el sistema. Asimismo, la selección del modelo adecuado es esencial para garantizar la exactitud en las predicciones, optimizando la efectividad del mantenimiento predictivo.

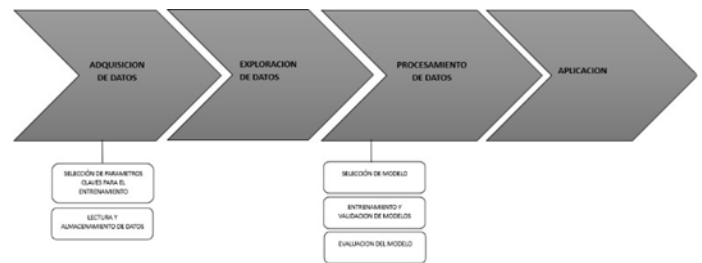


Fig. 2 Diagrama del proceso de Machine Learning

A. Adquisición de datos

1) Selección de parámetros clave para el entrenamiento:

En esta etapa, se definieron los parámetros más importantes que servirían de entrada para el modelo de machine learning con el fin de predecir cuándo la resistencia eléctrica alcanzará el valor crítico de 61 ohmios. Los parámetros considerados incluyen:

- Resistencia eléctrica (Ohmios): Variable objetivo del modelo.
- Tiempo (Fecha y hora): Parámetro independiente que describe la evolución temporal de la resistencia.

2) *Lectura y almacenamiento de datos:* La recopilación de datos se realizó de forma manual, utilizando un amperímetro para medir la resistencia eléctrica en intervalos de tiempo específicos. Este enfoque permitió garantizar la realidad y precisión de los valores registrados, asegurando que las mediciones reflejaran el comportamiento real del sistema. El proceso se llevó a cabo de la siguiente manera:

- Medición manual de valores: Se realizaron mediciones de resistencia eléctrica periódicamente con un amperímetro, anotando los resultados junto con la fecha y hora correspondientes.
- Organización de datos: Los valores recolectados se almacenaron en una base de datos estructurada para facilitar su análisis y posterior uso en el modelo de machine learning.
- Preparación de datos para el modelo: Se integraron los valores medidos como entradas para el entrenamiento y validación del modelo predictivo.

B. Exploración de datos

Se realizó un estudio preliminar de los datos recopilados con el propósito de identificar tendencias y validar su idoneidad para el modelo de machine learning. Las actividades incluyeron:

- Análisis temporal: Se verificó el comportamiento creciente de la resistencia a lo largo del tiempo.
- Identificación de patrones: Confirmación de que la resistencia presenta una progresión exponencial.

C) Procesamiento de datos.

1) *Selección de modelo:* La técnica seleccionada fue Regresión Lineal, debido a:

- Su simplicidad en términos de implementación y cálculo.

- La naturaleza exponencial del crecimiento de la resistencia eléctrica, que puede aproximarse linealmente a corto plazo.

2) *Entrenamiento y validación de modelos*: El modelo predictivo se entrenó y validó utilizando la información simulada, que fue segmentada de la siguiente manera:

- 80% de los datos: Utilizados para el proceso de entrenamiento.
- 20% de los datos: Empleados para validar el rendimiento del modelo.

Para evaluar el desempeño del modelo, se aplicaron métricas como:

- Validación cruzada (5 particiones)
- Baseline RMSE
- RMSE del modelo de regresión lineal
- MAPE

3) *Evaluación de modelo*: El rendimiento del modelo se evaluó mediante las métricas calculadas y un análisis detallado de los residuos.

D) *Aplicación Manual de Resultados en el HMI*

El ingreso manual de las fechas predichas representa un paso intermedio viable en el desarrollo del sistema. Permite validar el funcionamiento del HMI y las alertas visuales, al tiempo que facilita la adopción del mantenimiento predictivo en un entorno práctico y controlado.

III. SISTEMA SCADA (SUPERVISIÓN, CONTROL Y ADQUISICIÓN DE DATOS)

A. *Diseño*

La primera etapa de nuestro proyecto comienza con el diseño del sistema SCADA con el que nos ayuda a monitorear las variables que necesitamos para luego poder hacer el análisis con el modelo Machine Learning. Este sistema tiene como controlador un PLC 1200 del modelo 6ES7214- 1AG40-0XB0. Además, está conectado a un módulo de expansión para entradas analógicas del modelo 6ES7231- 4HD32-0XB0, dicho módulo sirve para las señales de entrada mandadas por los sensores. También está conectado por medio de ethernet a un SIMATIC HMI TP700 Comfort del modelo 6AV2124-0GC01-0AX0. Del mismo modo para la medición de voltaje se deberá usar el transmisor de voltaje con código HD-T1-3U, con entrada de 0 a 500 Vac y la salida para que sea leída por el plc es de 4-20mA DC. Del mismo modo para la medición de corriente se tendrá que usar el sensor SCT-013-000, con salida 4-20mA DC, la selección de este sensor es por la practicidad y por ser el más conocido para la medición de corriente en sistemas similares. Para la medición de temperatura se usaría el mismo sensor que está utilizando la máquina inyectora actualmente, que son las termocuplas tipo j. Este proceso de muestra en la Fig.3

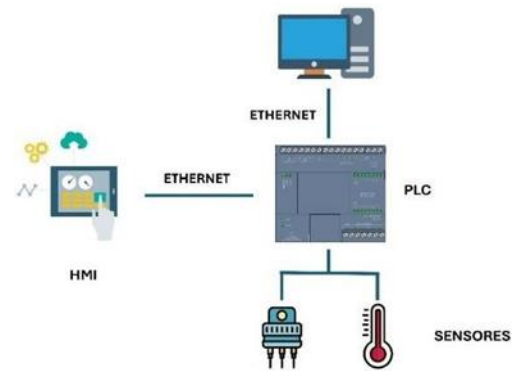


Fig. 3 Arquitectura SCADA

B. *Pantalla principal*

En la pantalla principal se visualiza las resistencias de acuerdo con la posición que se encuentran en la máquina para poder diferenciarlas de mejor manera y poder identificarlas rápidamente.

De la misma manera se muestran las variables por cada resistencia, como el voltaje, corriente, resistencia y temperatura. Además, se observa en la Fig.4a las opciones que tiene el Data Logger en el sistema, como la opción de ayuda y configuración de esta. Finalmente se muestra como activar la opción de grabar o no los datos en el sistema.

C. *Data logger en el sistema*

1) *Configuración del data logger en el SCADA:*

La data registrada sobre las variables se guardará dependiendo de la configuración dada mediante la pantalla en el HMI, dicha opción se muestra en la Fig.5a en la cual tenemos una opción de configuración para el data logger:

- Se puede seleccionar el registro a modificar
- Cambiar el nombre del archivo en el formato CSV.
- Modificar el periodo de muestras en segundos.
- También se puede borrar el data log seleccionado.
- Opción de comenzar a escribir los datos
- Se ubica led de muestreo de datos que parpadea de acuerdo con los periodos muestreados indicados.

2) *Descargar data logger en formato excel:*

Para poder obtener la data que ya hemos anteriormente grabado, debemos seguir los siguientes pasos para extraerla y tenerla como archivo csv.

3) *Conectar el plc a una computadora, en este caso se conecta mediante ethernet.*

4) *Ingresar al navegador de la computadora.*

5) *Escribir el IP de comunicación con el PLC e ingresar.*

6) *Ya en la interfaz de conexión con el PLC, ingresar a la cuenta admin.*

7) *Ingresar a la opción de navegador de archivos.*

8) Encontrar la carpeta data logger, y seleccionar el archivo que queremos descargar.



a)

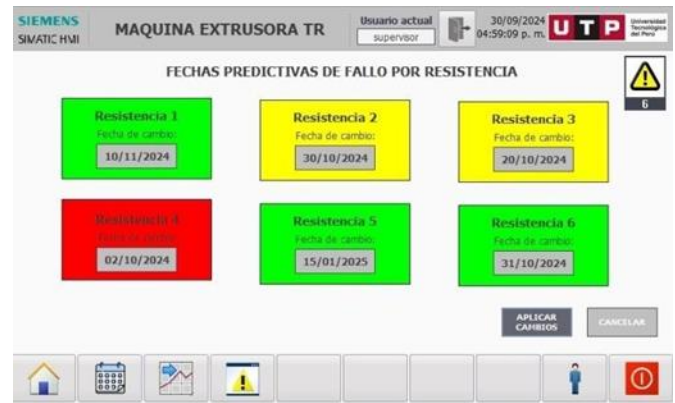


b)

Fig. 4 Pantallas HMI: a) Pantalla principal, b) Históricos



a)



b)

Fig. 5 Pantallas HMI: a) Configuración Data Logger, b) Fechas de predicción

D. Tendencias de variables en vivo

En la pantalla que se muestra en la Fig.4b, se puede apreciar las tendencias que tiene el sistema, de acuerdo con la resistencia que necesitamos ver, podemos seleccionarla en la parte izquierda de la pantalla, al seleccionar el número de resistencia, se podrá visualizar en la parte derecha un historial mediante una gráfica, que mostrará los valores durante el tiempo seleccionado de todas las variables como voltaje, corriente, ohmiaje, potencia y temperatura. Cabe resaltar que en la parte central se puede visualizar también los valores de las variables en una gráfica de tipo barras indicando numéricamente el valor exacto.

E. Fechas predictivas y alarmas gráficas

Las fechas predictivas de falla que salgan como resultado del análisis mediante el Machine Learning desarrollado, se pondrán manualmente en esta pantalla, donde hay un espacio derivado para cada resistencia eléctrica. También estos cuadros cambiarán de color de acuerdo con la aproximación a la fecha. El cambio de color tiene las siguientes condiciones:

- Si hay un intervalo más de un mes para llegar a la fecha indicada, el recuadro estará de color verde.
- Si hay un intervalo menor a un mes para llegar a la fecha indicada, el recuadro estará de color amarillo.
- Si hay un intervalo menor a una semana para llegar a la fecha indicada, el recuadro estará de color rojo.

La pantalla anteriormente descrita se encuentra en la Fig.5b

IV. MACHINE LEARNING

A. Base de datos

Para tener una base de datos realista se hizo las mediciones de todas las variables necesarias para hallar el ohmiaje, a partir del voltaje, la corriente y la temperatura. Se hizo este tipo de medición ya que la máquina es crítica

en la empresa, por lo que no puede parar la producción. Se eligió a la resistencia del cañón 1 de la estación 2, para hacer las mediciones durante dos meses aproximadamente. Dichas mediciones se realizaron mediante una pinza eléctrica y un multímetro de marca Fluke Durante la semana de lunes a viernes, en el horario de producción de 7:30 am a 5:30 pm.

1) Medición de voltaje:

Cada 30 minutos durante el turno de producción, de lunes a viernes, se hace la medición de voltaje con el multímetro Fluke, se hace la medición desde la caja de borneras, donde se encuentran las conexiones del relé de estado sólido a las resistencias. Se ubicó las borneras de la resistencia elegida para el análisis, y se hizo la marcación correspondiente para poder identificarla cada día en las mediciones diarias.

2) Medición de corriente:

Para esta medición se usó una pinza amperimétrica Fluke 376 FC, el cual tiene una herramienta que ayudó mucho en la medición, ya que usamos un celular para instalar la aplicación de Fluke Connect, esta aplicación nos permitió poder grabar las mediciones durante todo el turno, con la pinza amperimétrica colocada en la alimentación de la resistencia eléctrica que se analizó. Este tipo de herramienta nos facilitó en la exactitud de las mediciones y así poder tener los datos exactos. Diariamente se hizo la descarga de la data desde el celular al pc. Donde se filtró las mediciones de cada 30 minutos durante el turno, y de esa manera obtuvimos la base de datos de la corriente diariamente durante los dos meses en total.

3) Cálculo de ohmiaje:

Ya teniendo los datos de voltaje y corriente, simplemente en a base de datos se hizo el cálculo de ohmiaje con la ley de ohm y usamos la fórmula de $V/I=R$, obteniendo los ohmiajes durante los dos meses de la resistencia, y poder ver las variaciones que pudo tener durante ese tiempo.

B. Límite de ohmiaje

Se realizó la recolección de resistencias antiguas ya anteriormente retiradas de la máquina, se encontraban en un almacén de residuos viejos, de los cuales se seleccionaron 6 resistencias eléctricas para el análisis de voltaje, corriente y ohmiaje. Dichas resistencias dieron como resultado la medición con la alimentación de 223 voltios, 3.6 amperios y

61 ohmios, para llegar a estos datos se sacó un promedio de las 6 resistencias eléctricas que se seleccionó. Para acotar las resistencias tienen de potencia diseñada 1000 watts, al tener los datos mencionados antes según los cálculos solamente llegaría a los 800 watts aproximadamente, entonces la capacidad de la resistencia disminuyó en un 20%, y por esa razón se hacía los cambios, porque ya no llegaba a la temperatura que se desea en la máquina.

TABLA I
COMPARACIÓN DE PARÁMETROS ENTRE DISTINTOS ESTADOS EN LA RESISTENCIA

	VOLTAJE	CORRIENTE	OHMIAJE
RESISTENCIA DAÑADA	223 V	3.6 A	61 Ω
RESISTENCIA INSTALADA	223 V	4.4 A	50.7 Ω
RESISTENCIA NUEVA	223 V	4.5 A	49.5 Ω

C. Modelo de Machine Learning

Para elegir el modelo de Machine Learning más eficaz para el tipo de análisis se realizó una comparación entre los demás modelos, se llegó a la elección de regresión lineal, por el tipo de datos que se tienen de entrada que son el tiempo y el ohmiaje de la resistencia evaluada.

Se propone el uso de un modelo de regresión lineal para predecir la probabilidad de fallo en las resistencias eléctricas por las siguientes razones:

1) Relación entre tiempo y ohmiaje

La regresión lineal resulta adecuada cuando existe una relación proporcional o cercana a lineal entre el tiempo y el valor de ohmiaje. Si se observa que los datos muestran un patrón de incremento constante en el ohmiaje a medida que pasa el tiempo, este modelo permite capturar esa tendencia de forma efectiva. En sistemas eléctricos, como la degradación de componentes, los cambios suelen seguir un comportamiento lineal bajo condiciones estables [7].

2) Predicción de valores continuos

La tarea principal en este caso es estimar el valor del ohmiaje en función del tiempo, lo cual se alinea perfectamente con las capacidades de la regresión lineal. Este modelo es ideal para predecir valores numéricos continuos, además de ofrecer una estructura sencilla para interpretar cómo las variables afectan los resultados. Modelos como este han demostrado su eficiencia en problemas predictivos similares [8].

3) Determinación del momento crítico

Un beneficio clave de la regresión lineal es que, dado un umbral definido (61 ohmios en este caso),

permite calcular el tiempo exacto en que se alcanzará este valor. Esto se realiza resolviendo la ecuación de la recta obtenida tras entrenar el modelo. Este enfoque ha sido ampliamente empleado para identificar momentos críticos en sistemas físicos y predicción de fallas [9]

4) Simplicidad e interpretabilidad

La regresión lineal destaca por su simplicidad y claridad en la interpretación. Es un modelo matemáticamente directo que facilita la comprensión de los resultados y la influencia de las variables. Además, permite evaluar su precisión a través de métricas como el error cuadrático medio (RMSE) y el error absoluto porcentual medio (MAPE), lo que lo convierte en una herramienta confiable para análisis prácticos [10].

5) Validación y ajuste del modelo

El uso de este modelo es efectivo siempre que las condiciones básicas (linealidad, independencia de errores, homocedasticidad y normalidad de los residuos) se cumplan razonablemente. Además, el análisis gráfico y las métricas de desempeño ayudan a confirmar si el modelo ofrece resultados adecuados. Estas estrategias son comunes y efectivas en el ámbito del modelado predictivo [11].

6) Conclusión

La regresión lineal es un método sólido y práctico para este caso, ya que permite:

- Capturar relaciones lineales entre el tiempo y el ohmioaje.
- Realizar predicciones continuas de manera precisa.
- Identificar el momento exacto de falla basado en un umbral definido.
- Interpretar fácilmente los resultados y validar su calidad con análisis estadísticos.

D. Desarrollo de programación en google colab

Para desarrollar la programación de Machine Learning se usó google colab, una herramienta eficaz en este tipo de proyectos, el proceso del programa está dividido en 3 etapas y con el resultado final de visualización de resultados, estas etapas están divididas en subprocesos los cuales son los siguientes:

1) Definición de data:

- a) *Calcular timedays*: Determina el tiempo transcurrido desde el primer registro de datos.
- b) *Modelo de degradación*: Definir el modelo de degradación exponencial.
- c) *Parámetros iniciales*: Establecer R_0 y $failure_resistance$.

2) *Determinación de la función K*: La fórmula que nos ayudará para la predicción de fecha de falla, es la de crecimiento exponencial. Según Weisstein [12], el crecimiento exponencial describe procesos donde la tasa de cambio de una variable es directamente proporcional al valor de esta, generando un crecimiento acelerado. Dicha fórmula es la siguiente:

$$N(t) = N_0 e^{kt}$$

Donde:

- $N(t)$ es la cantidad en el tiempo t .
- N_0 es la cantidad inicial.
- k es la tasa de crecimiento.
- e es la base del logaritmo natural.

En este caso la fórmula se usará porque de acuerdo con el tiempo y a la tasa de crecimiento K , que se hallará en relación con las condiciones y valores de base de datos que se tiene, obtenemos una relación la cual indica un crecimiento en el ohmioaje durante el tiempo con relación a K . Y para hallar K se seguirán los siguientes pasos dentro de la programación de Machine Learning.

- a) *Función K*: Definir la función para hallar K .
- b) *Calcular K*: Ejecutar la función para encontrar el valor de K .
- c) *Proyección futura*: Calcular la proyección futura.
- d) *Día de falla*: Determinar el día de falla esperado.

3) Predicción de fecha de falla:

- a) *Día a fecha*: Cambiar día de falla a fecha específica.
- b) *Rango de fechas*: Crear un rango de fechas futuras.
- c) *Predicción de resistencia*: Predecir valores de resistencias futuras.

E. Entrenamiento y validación de modelo

Se utilizó un enfoque donde el 80% de los datos disponibles se destinó al entrenamiento del modelo, mientras que el 20% restante se reservó para validar su desempeño y asegurar la precisión de las predicciones.

F. Evaluación del modelo

Se emplearon las métricas MAPE, Baseline RMSE y RMSE de la regresión lineal para evaluar el modelo. El MAPE mide el error porcentual medio, facilitando una interpretación relativa al problema [13]. El Baseline RMSE se usó como referencia para comparar el modelo con un enfoque básico, evaluando si la regresión lineal representaba una mejora significativa [14]. Finalmente, el RMSE de la regresión lineal cuantifica la desviación

promedio entre predicciones y valores reales, siendo útil para datos con unidades específicas como el ohmiaje [15]. Estas métricas aseguran una evaluación integral del rendimiento del modelo.

G. Relación entre Predicciones y Residuos

Se elaboró un gráfico de dispersión para examinar cómo se distribuyen los residuos en función de las predicciones. Este análisis es fundamental para detectar posibles patrones en los errores y confirmar que el modelo cumple con sus supuestos [16].

V. RESULTADOS

A. Métricas del modelo

El desempeño del modelo de regresión lineal fue evaluado mediante métricas clave que demuestran su precisión y eficacia. En primer lugar, el MAPE (Error Porcentual Absoluto Medio) obtenido fue de 0.95%, lo que refleja un error promedio relativo muy bajo en comparación del ohmiaje predicho y el ohmiaje de la data validada. Esto indica que las predicciones del modelo están bastante alineadas con los valores reales, lo cual es fundamental en un escenario de predicción como este.

Por otro lado, el Baseline RMSE alcanzó un valor de 1.3021 ohmios, representando el error promedio al emplear un modelo base sin aprendizaje. En comparación, el RMSE del modelo de Regresión Lineal fue de 0.6407 ohmios, evidenciando una reducción significativa del error. Este resultado pone de manifiesto que el modelo fue capaz de identificar patrones en los datos y mejorar considerablemente la precisión de las predicciones frente a un enfoque más simple.

En conclusión, estas métricas confirman que el modelo es eficiente y confiable para la tarea de estimar el ohmiaje en función del tiempo, lo cual es esencial para predecir con exactitud la fecha en que se alcanzará el umbral de falla. Como resumen se presenta la siguiente tabla:

TABLA II
RESULTADOS DE MÉTRICAS

Métrica	Valor	Interpretación
MAPE	0.95%	El modelo predice los valores con un error promedio relativo muy bajo, menor al 1%.
Baseline RMSE	1.3021 ohmios	Representa el error de un modelo base sin aprendizaje.
Regresión Lineal RMSE	0.6407 ohmios	Indica que el modelo reduce significativamente el error en comparación con el baseline.

B. Validación cruzada

Se utilizó la validación cruzada con cinco particiones para evaluar la capacidad de generalización del modelo. En la Fig.6 se observa el Error Cuadrático Medio (RMSE) calculado para cada partición, obteniendo un promedio de 0.7280 ohmios. Aunque en la tercera partición el error es mayor, posiblemente debido a la variabilidad en los datos o a una distribución menos representativa, las demás particiones presentan errores similares al promedio. Esto demuestra que el modelo mantiene un rendimiento estable en diferentes subconjuntos del conjunto de datos, lo que respalda su eficacia predictiva.

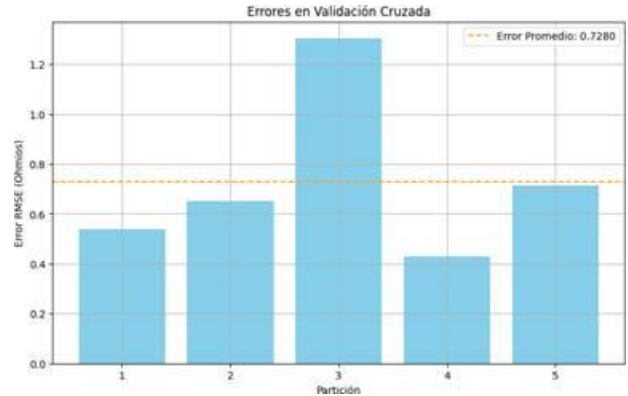


Fig. 6 Error en validación cruzada

C. Gráfico de Residuos vs Predicciones

La Fig. 7 permite evaluar el ajuste del modelo. Como se observa en la imagen, los residuos se distribuyen de forma uniforme alrededor de la línea cero (línea roja). Esto sugiere que el modelo no presenta patrones sistemáticos en los errores, lo cual indica un buen ajuste.

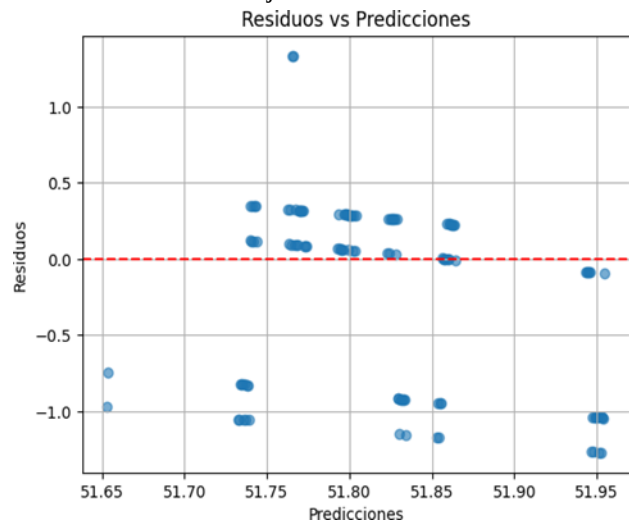


Fig. 7 Residuos vs Predicciones

D. Proyección y Validación de la Falla

El gráfico de proyección muestra cómo el modelo predice que la resistencia alcanzará el umbral crítico de 61 ohmios el 04 de julio de 2025. El área sombreada representa el intervalo de confianza, indicando la certeza del modelo sobre las predicciones. Los puntos validados confirman que el modelo logra un buen ajuste a los datos observados, proporcionando información precisa para el mantenimiento predictivo.

La Fig. 8 muestra la proyección y validación de la resistencia eléctrica R3 a lo largo del tiempo, con el objetivo de predecir la fecha exacta en la que alcanzará el umbral de falla definido en 61 ohmios.

- 1) *Resistencia Observada (línea azul sólida):* Representa los datos medidos de resistencia en el tiempo hasta el momento actual.
- 2) *Proyección del Modelo (línea naranja discontinua):* Muestra cómo el modelo de regresión lineal predice que la resistencia continuará aumentando en función de la tendencia observada en los datos históricos.
- 3) *Intervalo de Confianza (área sombreada naranja):* Indica el rango dentro del cual es probable que se encuentren los valores futuros de resistencia. Este intervalo refleja la incertidumbre en las predicciones del modelo.
- 4) *Predicción en Validación (puntos rojos):* Estos puntos son mediciones validadas del modelo, lo que indica que las predicciones se ajustan razonablemente bien a los datos reales en este rango.
- 5) *Umbral de Falla (línea roja horizontal):* Representa el valor crítico de 61 ohmios, que indica que la resistencia ha fallado.
- 6) *Fecha Estimada de Falla (línea verde discontinua):* Marca la fecha proyectada en la que el modelo predice que la resistencia alcanzará el umbral de falla. Según el gráfico, la fecha estimada de falla es el 04 de julio de 2025.



Fig. 8 Proyección y validación de la falla en resistencia R3

VI. CONCLUSIONES

- El sistema SCADA desarrollado demuestra que es posible integrar funcionalidades de predicción y monitoreo manual mediante un diseño eficiente en TIA Portal, con 6 pantallas de HMI y entradas simuladas para probar el sistema. Este enfoque permite una operación intuitiva y flexible para los operadores.
- La implementación del sistema de alertas por identificación de colores (verde, amarillo, rojo) facilita una rápida interpretación del estado de las resistencias por parte de los operadores. Este método es efectivo para priorizar acciones de mantenimiento basándose en la cercanía de las fallas predichas.
- La validación cruzada, combinada con métricas como MAPE (0.95%), Baseline RMSE (1.3021) y Regresión Lineal RMSE (0.6407), demuestra que el

modelo de machine learning tiene un desempeño altamente preciso, lo que refuerza su utilidad para el mantenimiento predictivo.

VII. RECOMENDACIONES

- El sistema SCADA podría automatizarse conectándolo directamente al modelo de Machine Learning, y tener una alarma directa, y ya no estar poniendo las fechas manualmente en el HMI.
- Se puede aumentar una alarma externa de sonido o ventana emergente para que aparezca cuando ya esté muy cerca la fecha de fallo de alguna resistencia. Eso permitiría dar prioridad a cambio y un mejor aviso.

- Se podría probar el Machine Learning con una base de datos con el historial de inicio de puesta en marcha de la resistencia a fin de fallo de esta.

REFERENCIAS

- [1] “Partes de una Inyectora de Plástico| Aeromaquinados”. Aeromaquinados. Disponible: <https://aeromaquinados.com/partes-de-una-inyectora-de-plastico/>
- [2] “Resistencias eléctricas para inyección de materiales - Resistencias Eléctricas Industriales”. Resistencias Eléctricas Industriales. Disponible: <https://juliomartineznaya.com/resistencias-electricas-para-inyectoras/>
- [3] Computer Simulation Experts, “Mantenimiento predictivo: la prevención de problemas como clave para industrias más sostenibles y competitivas,” ESSS, 06-Oct-2023. [Online]. Available: <https://www.esss.com/es/blog/mantenimiento-predictivo/>. [Accessed: 03-Nov-2024].
- [4] J. Lee, H.-A. Kao, and S. Yang, “Service innovation and smart analytics for industry 4.0 and big data environment,” *Procedia CIRP*, vol. 16, pp. 3–8, 2014.
- [5] “Libro Organización Y Gestión Del Mantenimiento,” PMM Business School | Escuela Operativa Internacional, 04-Aug-2021. [Online]. Available: <https://pmm-bs.com/libro-organizacion-y-gestion-del-mantenimiento-mantenimiento-como-negocio-balanced-scorecard/>. [Accessed: 03-Nov-2024].
- [6] F. Sáenz, “Paradas no planificadas. Causas y soluciones,” *Fernandosaez.com*, 25-Apr-2022. [Online]. Available: <https://www.fernandosaez.com/paradas-no-planificadas-causas-y-soluciones/>. [Accessed: 03-Nov-2024].
- [7] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*, 2nd ed. Springer, 2021.
- [8] A. Montgomery and K. Srinivasan, “Regression Analysis and Its Application in Predictive Modeling,” *Journal of Forecasting*, vol. 35, no. 4, pp. 415–430, 2016.
- [9] J. P. Casey, “Predictive maintenance using regression analysis,” *IEEE Transactions on Reliability*, vol. 65, no. 2, pp. 610–620, Jun. 2016.
- [10] D. Freedman, “Statistical models: theory and practice,” Cambridge University Press, 2nd ed., 2009.
- [11] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, 2009.
- [12] E. W. Weisstein, “Exponential Growth,” *MathWorld—A Wolfram Web Resource*, Wolfram Research, 2002.
- [13] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2011.
- [14] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*, 2nd ed. Springer, 2021.
- [15] C. Bishop, *Pattern Recognition and Machine Learning*, 1st ed. Springer, 2006.
- [16] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*, 2nd ed. Springer, 2021.