

Detección de parámetros críticos en procesos de pintura automotriz con enfoque hacia industria 4.0

Edmundo Álvarez Solís¹, Isaías Emmanuel Garduño², Juan Roberto Moreno García^{3,4}, Javier Méndez Lozoya^{3,5},
y Fernando Sebastián Chiwo^{3,6,*}

Posgrado CIATEQ A.C.¹, CONAHACYT—CIATEQ A.C.², Universidad Marista de San Luis Potosí³, Universidad Tecnológica de San Luis Potosí⁴,
Instituto Tecnológico Nacional de México⁵, CIATEQ A.C.⁶

Aguascalientes, Ags.¹; San Luis Potosí, S.L.P.^{2,3,6}; Soledad de Graciano Sánchez, S.L.P.^{4,5}; México

*Autor de correspondencia: fernando.chiwo@ciateq.mx

Abstract— In the automotive industry, the painting process is made up of several sub-processes that require meticulous control of functional parameters to ensure the production of high-quality parts. These subprocesses, although already established and standardized by multiple operating routines, require optimization to provide a higher-quality final product. This work proposes a data-based approach and the implementation of different machine learning algorithms to predict the behavior of different critical parameters within the automotive painting process. The results obtained are promising because they present low error percentages due to the optimal characteristics of the data used, such as value, volume, validity, and variability. This research contributes to improving predictive capabilities and optimizing the automotive painting process to improve efficiency and quality. This highlights the need to comply with established standards and parameters and the crucial role of information in decision-making, especially in the era of Industry 4.0, where digitalization and connectivity facilitate the implementation of machine learning models as a fundamental tool to improve process management and automate complex tasks.

Keywords— *Industry 4.0, Artificial Intelligence, Big data, Machine Learning, Forecast, Automotive, Paint shop.*

Resumen— Dentro de la industria automotriz, el proceso de pintura se compone de varios subprocesos que requieren un control meticuloso de los parámetros funcionales para garantizar la producción de piezas de alta calidad. Dichos subprocesos, que aunque ya están establecidos y estandarizados por múltiples rutinas de operación, requieren ser optimizados para proporcionar un producto final de mayor calidad. En este trabajo se propone una aproximación basada en datos y la implementación de diferentes algoritmos de aprendizaje automático para pronosticar el comportamiento de diferentes parámetros críticos dentro del proceso de pintura automotriz. Los resultados obtenidos son prometedores debido a que presentan porcentajes de error bajos debido a las características óptimas de los datos utilizados, como son el valor, el volumen, la validez y la variabilidad. Esta investigación contribuye a mejorar las capacidades predictivas y optimizar el proceso de pintura de automóviles para mejorar la eficiencia y la calidad. En esta se resalta la necesidad de cumplir con normas y parámetros establecidos y el papel crucial de la información en la toma de decisiones, especialmente en la era de la industria 4.0, donde la digitalización y conectividad facilitan la implementación de modelos de aprendizaje automático como una herramienta fundamental para mejorar la gestión de los procesos y automatizar tareas complejas.

Palabras claves— *Industria 4.0, Inteligencia Artificial, Big data, Modelos de aprendizaje automático, Pronósticos, Automotriz, Planta pintura.*

I. INTRODUCCIÓN

El entorno competitivo global en diversas industrias, especialmente en la automotriz, demanda la adopción de nuevas herramientas, metodologías y tecnologías. En este contexto, el concepto de la Industria 4.0 implica la incorporación de herramientas inteligentes de producción, con el objetivo de incrementar la eficiencia y productividad en diferentes procesos de manufactura, siempre y cuando, se garantice la satisfacción de los consumidores (Saquicela, 2020). La implementación de herramientas y recursos de la Inteligencia Artificial y el análisis de grandes volúmenes de datos, conocido como Big Data, permite la obtención de una cantidad virtualmente ilimitada de datos, con lo cual es posible realizar una caracterización óptima de los procesos de interés, que permitan la toma de decisiones fundamentadas

gracias a la abundancia de información. Esto implica la eliminación de desperdicios que no agreguen valor al producto y la reducción de la variabilidad (Younas, 2019; Dobos *et al.*, 2018; Mahesh, 2020; Carleo *et al.*, 2019).

Los tres propósitos principales del proceso de pintura automotriz son: evitar la corrosión, garantizar la apariencia y la durabilidad de la pintura, con esto se busca disminuir los tiempos de respuesta ante las anomalías y reducir los riesgos asociados al producto. El proceso se conforma por diferentes etapas, las cuales son “fosfatizado”, “electrodeposición”, “sello” y “esmalte”. En la presente investigación, se considera la etapa de esmalte, la cual consiste en la aplicación de tres capas de pintura, de forma manual por operadores y automática por robots. La primera capa de pintura, denominada como “pintura de fondo” o “primer”, cuyo principal objetivo es proteger la pintura de la radiación solar, es decir, garantizar la durabilidad de las capas de pintura contra los rayos ultravioleta. La segunda capa es conocida como “base”, y su función es proporcionar el color que percibe el cliente. Finalmente, la tercera capa aplicada es conocida como “barniz”, la cual otorga el brillo a la carrocería, y dureza para garantizar protección y durabilidad en todas las capas de pintura. Adicionalmente a la etapa de esmalte, es necesario aplicar un subproceso de horneado, con el cual las capas de pintura aplicadas en esmalte son deshidratadas y posteriormente secadas. La base de datos de las etapas esmalte y horneado la constituye múltiples lecturas de 10 parámetros críticos, evaluados en el 100% de los automóviles procesados. Al finalizar la etapa de adquisición de datos, se obtiene un promedio de 750 automóviles por día.

La finalidad de evaluar los parámetros críticos y garantizar el cumplimiento de sus respectivas normas de forma independiente, es que la variación anormal observada esté estrechamente relacionada con defectos en las unidades (Mazur, 2019). Entre los defectos observados, se encuentran el acumulamiento de material que escurre por efectos de la humedad y temperatura, y la pintura hervida, que resulta en la formación de burbujas sobre la superficie. Dichos defectos impactan negativamente en los tres propósitos principales del proceso de pintura automotriz. Estos aspectos contribuirán a optimizar los costos de manufactura y ocupación del mercado, implementando modelos de aprendizaje automático con pronósticos, de los diferentes parámetros en el proceso de pintura automotriz, y así gestionar las actividades de contramedidas y ajustarlas mediante el departamento responsable dichas actividades evitando defectos en las unidades implicando desarrollar herramientas estadísticas que posibiliten la inferencia y la predicción de parámetros alejados de las normas (Mazur, 2019).

En la etapa estudiada el flujo promedio de automóviles en proceso es de 90, equivalente a tres horas de producción, hasta llegar al primer punto de inspección, que es donde se detonan las actividades correctivas. En esta línea de producción, se verifica el cumplimiento de las normas respectivas, de los 10 parámetros en cuestión, pero siempre las actividades realizadas son correctivas, es decir, hasta que se presenta el defecto. Anteriormente, las técnicas utilizadas para realizar pronósticos eran el cálculo de la media aritmética y la ecuación de la recta, la primera con una variación promedio del 26% y la segunda con una variación promedio 17%, ambos métricos como margen de error contra el valor real. La parte central de esta investigación consiste en prever las variaciones de los 10 con un margen de error deseado menor al 10% respecto al valor real. En la época actual, existen una gran variedad de modelos y técnicas para realizar predicciones o pronósticos mediante algoritmos de Machine Learning o Aprendizaje Automático (Anderson Damián Jiménez Alfaro, 2021). Dichos algoritmos se dividen principalmente en dos tipos, de acuerdo a su funcionalidad, los algoritmos de aprendizaje supervisado y los algoritmos de aprendizaje no supervisado. En el primero, se entrena un algoritmo tomando en cuenta las características, las variables de respuestas, o bien, las etiquetas del proceso en el cuál se está trabajando, mientras que en el segundo, se entrena un algoritmo sin información previa acerca del proceso. Dentro del aprendizaje supervisado, existen dos tareas principales: la clasificación y la regresión, las cuáles se pueden adaptar

para realizar predicciones. En general, los algoritmos de aprendizaje supervisado se componen de dos fases: la fase de entrenamiento y la fase de prueba, destinando un 70-80% del conjunto de los datos para el entrenamiento, y un 20-30% para la fase de prueba (Mahesh, 2020).

El presente trabajo de investigación, propone un esquema de trabajo basado en el análisis de datos, que permita reducir la complejidad y la interconexión dentro de las industrias que presentan desafíos nuevos que pueden ser abordados por la automatización basada en el análisis de cantidades masivas de datos. Un aspecto esencial para la correcta implementación de algoritmos de aprendizaje automático es la adquisición de datos que se lleva a cabo por medio de la conectividad entre equipos, instrumentos y controladores lógicos programables en un proceso de manufactura, de esta manera la base de datos se construye con información de calidad (Alfaro & Ospina, 2021; Kebisek et al., 2020).

II. MATERIALES Y METODOS

En esta investigación, se implementaron diferentes modelos de aprendizaje automático supervisado para realizar pronósticos de los posibles valores que adquieren los diferentes parámetros, evaluados a lo largo del complejo proceso de pintura automotriz, y predecir cuándo es posible que los valores estén fuera de las especificaciones técnicas. Los algoritmos implementados fueron la regresión lineal, la regresión de procesos gaussianos, máquinas de soporte vectorial, y las redes neuronales. El modelo de regresión lineal se utiliza para determinar la existencia de relación lineal entre las variables estudiadas, presentada en la ecuación (1).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon \quad (1)$$

Donde y es variable dependiente, x_i es la variable independiente, β_0 indica el valor constante de la ecuación de regresión, $\beta_m x_m$ indica el coeficiente de regresión asociado a la variable independiente, ε expresa el error del modelo y m indica el número total de variables del modelo. Para las redes neuronales, el modelo característico se expresa mediante la ecuación (2):

$$a = f(wp + b) \quad (2)$$

El valor de salida se expresa con la variable a , f indica la función de transferencia, w es el peso escalar, p indica la entrada escalar y b expresa el sesgo escalar. Para la regresión de proceso gaussiano, el modelo se expresa mediante la ecuación (3):

$$y = x^T \beta + \varepsilon \quad (3)$$

El valor de salida o respuesta se expresa con la variable y , x indica el vector de entrada, ε el error de la varianza, β el coeficiente de estimación, T el espacio dimensional. Para la máquina de soporte vectorial, el modelo se expresa mediante la ecuación (4):

$$f(z): |y_n - (x_n' \beta + b)| \leq \varepsilon \tag{4}$$

Adicionalmente a estos modelos, también se consideró la raíz del error cuadrático medio (RMSE, por sus siglas en inglés Root Mean Squared Error) como estadístico de prueba para evaluar el rendimiento de los modelos implementados. Este parámetro, se expresa por medio de la ecuación (5):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \tag{5}$$

Para la cual n indica el tamaño de la muestra, P indica el valor observado, O es el valor esperado, e i que indica la lectura en turno. Todos algoritmos fueron implementación utilizando el módulo de Machine Learning del software MATLAB versión 2023b de The MathWorks (MATLAB, 2023). En la Fig. 2 se muestran los pasos a seguir para utilizar la paquetería indicada.

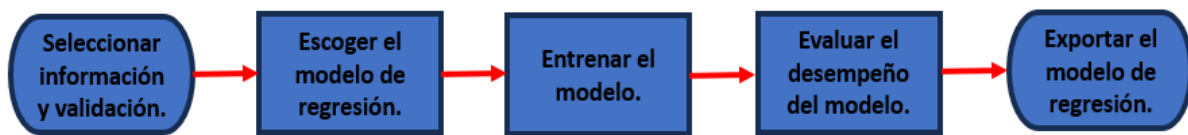


Fig. 2. Pasos para en el entrenamiento y validación de un modelo de aprendizaje automático.

Para la organización de los datos, se tomaron 10 parámetros críticos, y para cada uno de ellos se implementó un modelo. Cada parámetro, en su modelo de aprendizaje, es seleccionado como variable de respuesta, y se busca validar qué valor va a tomar en función de los valores de las variables predictoras, que son los 9 parámetros restantes. El etiquetado de los parámetros de control se muestran en la Tabla I, junto a sus respectivas unidades de medición.

Tabla I. Descripción de parámetros utilizados y unidad de medida.

Parámetro	Unidad de medida
Temperatura base 1	°C
Temperatura base 2	°C
Temperatura barniz	°C
Temperatura horno zona 1	°C
Temperatura horno zona 2	°C
Temperatura horno zona 3	°C
Temperatura horno heat up	°C
% humedad base 1	g/m ³
% humedad base 2	g/m ³
% humedad barniz	g/m ³

La información de los parámetros se recopila mediante controladores lógicos programables en tiempo real para todas las unidades producidas. La Figura 1 muestra la distribución de los sensores por las diferentes secciones de la etapa de esmalte, señalizados con puntos rojos y círculos negros en el exterior,

al igual que tablas que indican el nombre del parámetro evaluado por cada sensor, así como su unidad de medida.

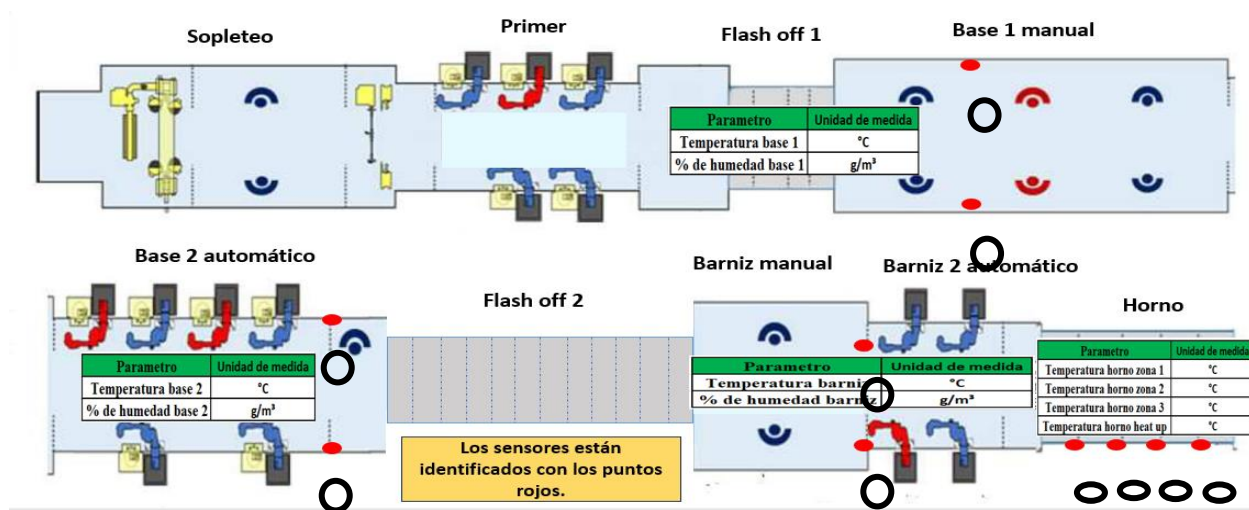


Fig. 1. Esquema de la cabina de esmalte ubicando los sensores (puntos rojos) con sus respectivas variables de medición.

III. RESULTADOS

Después de haber definido la estructura de la información, se utilizaron 1200 lecturas por parámetro. Es importante mencionar que cada lectura equivale a un automóvil procesado, con un promedio de 32 unidades por hora. Esto significa que para cada parámetro se registran 32 lecturas en una hora. Para la etapa de entrenamiento de las 1200 lecturas, se utilizó el 70% de los datos, dejando el 30% restante para la validación, lo que equivale a 840 y 340 lecturas respectivamente. Es decir, para cada modelo entrenado se utilizaron 840 lecturas en la variable dependiente y 7560 en las variables independientes. Para la etapa de validación se emplearon 340 lecturas en la variable dependiente y 3060 en las variables independientes.

Para evaluar el rendimiento de los modelos de regresión en la etapa de entrenamiento, se considera el RMSE como métrica de evaluación, considerando aquel modelo que presente un valor menor, como el más óptimo, ya que es con el que se obtiene un menor error. La misma métrica se aplica en la etapa de validación para cada modelo. Los modelos de aprendizaje automático son óptimos cuando presentan un 90% de resultados verdaderos positivos en las etapas de entrenamiento y validación, es decir, cuando el RMSE es del 10% o menos.

Como se puede observar para los diez parámetros evaluados, el algoritmo de regresión lineal presentó el menor porcentaje de error. Se proporciona el estado de la etapa de entrenamiento y validación, así como el porcentaje de error para cada una de las etapas, como se muestra en la Tabla II.

Tabla II. Parámetros y modelos de regresión entrenados.

Parámetro	Tipo de modelo	Entrenamiento	Validación
		RMSE	RMSE
% humedad base 1	Gaussian Process Regression	4.12 %	3.98 %
% humedad base 2	Gaussian Process Regression	3.68 %	3.70 %
Temperatura horno heat up	Gaussian Process Regression	6.64 %	6.96 %
Temperatura horno zona 1	Neural Network	5.57 %	5.28 %
Temperatura horno zona 2	Gaussian Process Regression	3.73 %	3.71 %
Temperatura horno zona 3	Linear Regression	3.94 %	4.14 %
% humedad barniz	Support Vector Machine	3.46 %	3.15 %
Temperatura barniz	Gaussian Process Regression	3.16 %	3.37 %
Temperatura base 2	Gaussian Process Regression	3.24 %	3.24 %
Temperatura base 1	Gaussian Process Regression	3.29 %	3.22 %

Para el parámetro "Temperatura barniz", el valor (RMSE) en la etapa de entrenamiento fue de 3.17% y en la etapa de validación fue de 3.38%, lo cual es un resultado óptimo debido a que el porcentaje de error fue inferior al 10%. El modelo de regresión que presentó mejor ajuste fue el proceso de regresión Gaussiana. Dentro de los 4 modelos de aprendizaje entrenados y validados, este parámetro es el que muestra el resultado más eficiente, al obtener el menor porcentaje de error entre el pronóstico y el valor real. En la etapa de entrenamiento, visualmente, como se observa en la Figura 3, está conformada por tres gráficas identificadas por incisos.

En la Figura 3A, se aprecia el ajuste de la variable de respuesta señalizada con puntos amarillos, superpuesta contra el modelo de aprendizaje automático con puntos azules, lo cual nos indica que existe una relación entre las variables estudiadas. En la Figura 3B, se observa el ajuste del modelo de predicción contra el valor actual señalizado con puntos azules, tomando como referencia el ajuste de la pendiente de la recta, lo que indica que existe una relación entre las variables estudiadas en la predicción del valor buscado. En la Figura 3C se muestran los residuos de los errores del modelo de aprendizaje automático. Se observa que los valores se distribuyen por encima y por debajo de la recta que indica el valor "0", lo que significa aleatoriedad en los residuos de los errores.

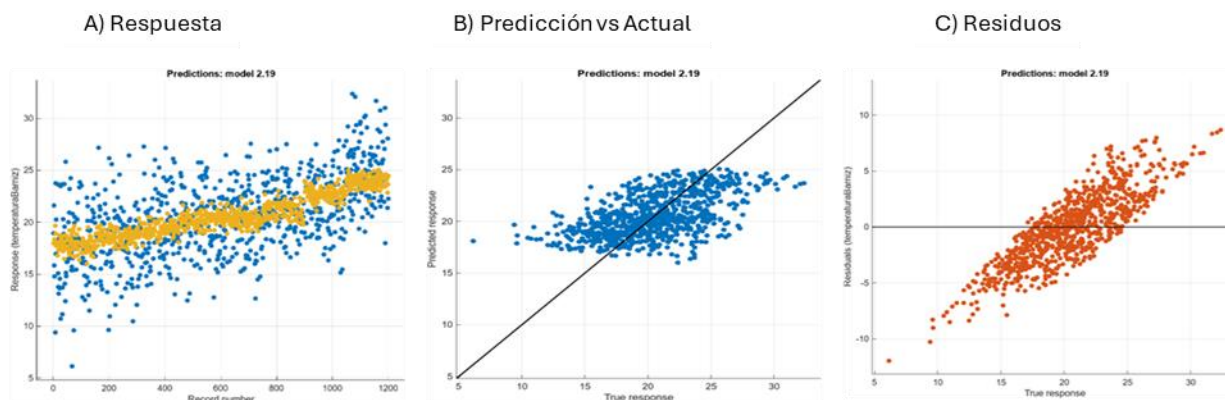


Fig. 3. Gráficos del modelo de aprendizaje automático para el parámetro “Temperatura de barniz”: A) Se observa una tendencia lineal en las respuestas predecidas (color amarillo) en comparación con las respuestas leídas (color azul), B) el ajuste del modelo muestra una aglomeración en un rango de 15 a 25 °C, con una variabilidad producida por la gran cantidad de datos analizada, C) los residuos entre las lecturas reales y las lecturas predecidas muestra una tendencia lineal, donde se confirma el rango de 15 a 25 °C como el óptimo para esta variable de respuesta.

En la etapa de validación, se realiza un análisis gráfico similar al de la etapa de entrenamiento como se observa en la Figura 4. En la Figura 4A, se observa el ajuste del modelo de predicción contra el valor actual, señalizado con puntos azules, tomando como referencia el ajuste de la pendiente de la recta, lo que indica que existe relación entre las variables estudiadas en la predicción del valor buscado. En la Figura 4B, se muestran los residuos de los errores del modelo de aprendizaje automático. Se observa que los valores se distribuyen arriba y debajo de la recta que indica el valor del "0", lo que significa aleatoriedad en los residuos de los errores.

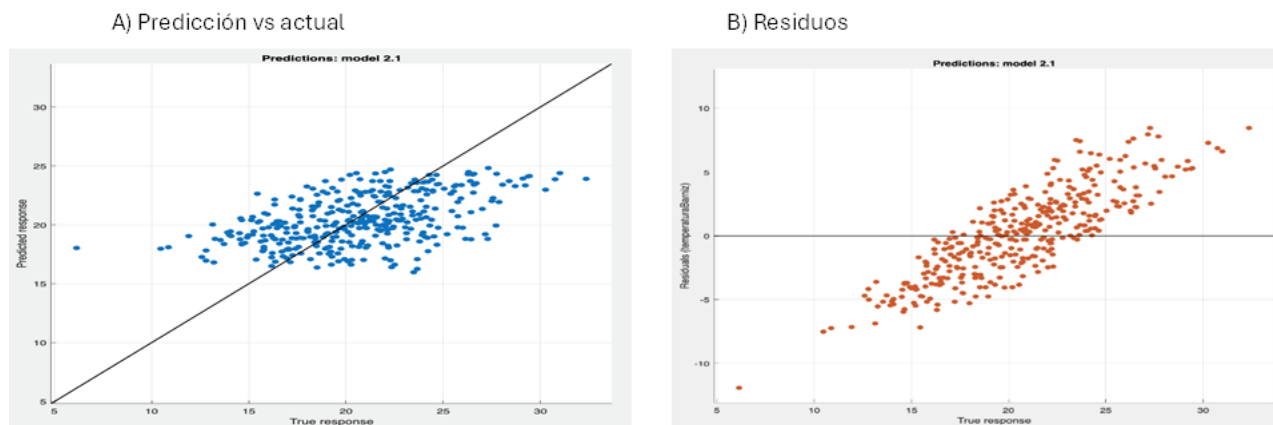


Fig. 4. Modelo del ajuste de regresión para la etapa de validación, donde se utilizó un 30% del conjunto total de los datos: A) el ajuste sigue la misma tendencia observada en la etapa de entrenamiento, B) la gráfica de residuos, de igual forma, indica la misma tendencia, indicando así que el modelo es óptimo para esta variable de respuesta.

Como se puede observar, en todos los modelos de aprendizaje desarrollados, el valor de RMSE es menor al 10%, lo que indica un desempeño más eficiente que los métodos utilizados anteriormente, como la media aritmética y la ecuación de la recta. La implementación de estos 10 modelos de aprendizaje automático, desarrollados a partir de la información de los 10 parámetros críticos del proceso de pintura automotriz en la etapa de esmalte, permite pronosticar de manera oportuna los valores que pueden tomar los parámetros a lo largo del tiempo. Esto facilita la realización de contramedidas preventivas para ajustarlos, evitando así impactar en la calidad de las unidades debido a valores de parámetros fuera de norma.

Después de desarrollar estos 10 modelos, se creó un macro en un software de hoja de cálculo vinculado a la base de datos. Dicho macro registra la información de los parámetros y, mediante fórmulas, carga el modelo de regresión correspondiente a cada uno. De esta manera, automáticamente cada 5 minutos se realizan pronósticos para validar el cumplimiento de las normas. Anteriormente, la confirmación de los parámetros solo se revisaba dos veces por turno y se registraba en una hoja de verificación. Luego, la frecuencia de confirmación se incrementó exponencialmente.

IV. CONCLUSIÓN

El presente trabajo realizado conforma un análisis basado en datos sobre los aspectos críticos en los procesos de pintura automotriz, en el contexto de la industria 4.0. Destaca la importancia del proceso de pintura en la manufactura automotriz, evidenciando significativamente su aporte a los costos totales de operación y la necesidad crítica de cumplir con normas de los parámetros establecidos para garantizar la calidad del producto final. Además, se subraya la importancia de la información para entornos competitivos, enfatizando cómo la toma de decisiones informadas puede reducir la incertidumbre y minimizar errores. Esta reflexión se relaciona con la transformación hacia la industria 4.0 y la aceleración digital, resaltando la interconexión de componentes y la recolección de datos para un análisis más eficiente y proactivo. Finalmente, se recalca el papel vital de los modelos de aprendizaje automático para mejorar la gestión y la toma de decisiones, ofreciendo una herramienta poderosa para prevenir errores y optimizar el cumplimiento de normas, reduciendo significativamente el margen de error en los pronósticos de los parámetros durante la producción. En conclusión, es importante recalcar la necesidad imperante de optimizar procesos dentro de la industria automotriz, no solo para asegurar la calidad del producto final, sino también para potenciar la eficiencia y competitividad en un entorno industrial cada vez más dinámico y exigente.

REFERENCIAS

- Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L., & Zdeborová, L. (2019). Machine learning and the physical sciences. *Reviews of Modern Physics*, 91(4). <https://doi.org/10.1103/revmodphys.91.045002>
- Dobos, P., Tamás, P., Illés, B., & Balogh, R. (2018). Application possibilities of the Big Data concept in Industry 4.0. *IOP conference series. Materials science and engineering*, 448, 012011. <https://doi.org/10.1088/1757-899x/448/1/012011>
- Grant, J. (2021). *Aprendizaje Automático Profundo: Aprenda inteligencia artificial, algoritmos de máquina utilizando técnicas y métodos avanzados de aprendizaje profundo de máquinas*. Independently Published.

- Kebisek, M., Tanuska, P., Spendla, L., Kotianova, J., & Strelec, P. (2020). Artificial intelligence platform proposal for paint structure quality prediction within the industry 4.0 concept. *IFAC-PapersOnLine*, 53(2), 11168–11174. <https://doi.org/10.1016/j.ifacol.2020.12.299>
- Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*, 9(1), 381–386.
- MatLab. (2023). version R2023a. Natick, Massachusetts: The MathWorks Inc.
- Mazur, M. (2019). Quality assurance processes in series production of car elements. *Quality Production Improvement - QPI*, 1(1), 610–617. <https://doi.org/10.2478/cqpi-2019-0082>
- McCalley, B. W. (2012). *Model T ford: The car that changed the world*. Krause Publications.
- Rai, R., Tiwari, M. K., Ivanov, D., & Dolgui, A. (2021). Machine learning in manufacturing and industry 4.0 applications. *International Journal of Production Research*, 59(16), 4773–4778. <https://doi.org/10.1080/00207543.2021.1956675>
- Saquicela, J. L. S. (2020). Transformación digital de la industria 4.0. *Polo del Conocimiento: Revista científico-profesional*, 5, 1344–1356.
- Younas, M. (2019). Research challenges of big data. *Service Oriented Computing and Applications*, 13(2), 105–107. <https://doi.org/10.1007/s11761-019-00265-x>