

DetECCIÓN DE FALLAS MECÁNICAS MEDIANTE "MACHINE LEARNING", UTILIZANDO EL CLASIFICADOR "RANDOM FOREST"

Daniel Junio Soares Rodrigues¹, Gabriel do Carmo Dantas²,
Geraldo Roberto de Sousa³, Lucas Costa Brito⁴, Jorge Nei Brito⁵

¹Gep_LASID, DEMEP, UFSJ, Brasil. danieljunio432@yahoo.com.br

²Gep LEREL, DCNAT, UFSJ, Brasil. dan.tas@hotmail.com

³Gep_LASID, DEMEP, UFSJ, Brasil. geraldor@ufsj.edu.br

⁴LAV, FEMEC, UFU. brito.lcb@gmail.com

⁵Gep_LASID, DEMEP, UFSJ, Brasil. brito@ufsj.edu.br

Resumen

En este trabajo se presentará una Inteligencia Artificial (IA) para el seguimiento de defectos de origen mecánico (desequilibrio, desalineación y holgura mecánica), además de la condición sin defectos, a partir de señales de vibración. El uso de "Machine Learning" puede considerarse un instrumento dentro de la Inteligencia Artificial para el diagnóstico de fallas mecánicas en máquinas rotativas. En este trabajo se abordó el modelo de "Machine Learning", a través de "Supervised Machine Learning" y utilizando "Classification Method" con algoritmos de clasificación por "Random Forest". El modelo se entrenó con el 70% de los datos disponibles en la "Base de Datos" y el 30% se utilizó para la validación de la prueba. Para la condición "Sin Defecto", el porcentaje de precisión fue del 99% y para las excitaciones "Desequilibrio" y "Holgura Mecánica", fueron del 98,1% y 99,3%, respectivamente. Esto se debe a la gran cantidad de señales disponibles para la prueba (532, 283 y 557) respectivamente. Para "Desalineamiento", el porcentaje de respuestas correctas fue del 69,6%. También fue influenciado por los porcentajes de error de 11,2% y 19,2%, por "Desequilibrio" y "Holgura Mecánica", respectivamente. Esto se debe al bajo número de señales disponibles para la prueba, solo 28, y porque las excitaciones están relacionadas con frecuencia de rotación (f_r) y sus armónicos, lo que provoca "Confusión". De manera similar al razonamiento anterior, "Holgura Mecánica + Desalineación" presentó el porcentaje de acierto del 46,9%, con una contribución de error del 40,4% por "Desequilibrio". Esto también se debe a las mismas razones que antes, siendo aún menor el número de señales disponibles para las pruebas, es decir, solo 18. Para aumentar el asertividad, es necesario tener un mayor número de señales, o aplicar el "Data Augmentation" para aumentar el número de señales o incluso extraer más parámetros discriminativos en el modelo. Los resultados muestran que la metodología propuesta permitió la detección de fallas supervisadas en máquinas rotativas, siendo una herramienta prometedora para ser aplicada en la Industria 4.0.

Palabras clave: Detección de Fallas, "Machine Learning", "Random Forest", Inteligencia Artificial.

Abstract

An Artificial Intelligence (AI) will be presented for monitoring defects of mechanical origin (imbalance, misalignment, and mechanical backlash), in addition to the condition without defect, from vibration signals. The use of "Machine Learning" can be considered an instrument within Artificial Intelligence for the diagnosis of mechanical failures in rotating machines. In this work, the "Machine Learning" model was approached, through "Supervised Machine Learning" and using "Classification Method" with classification algorithms by "Random Forest". The model was trained with 70% of the data available in the "Database" and 30% were used for test validation. For the "No Defect" condition, the percentage of success was 99% and for the "Unbalance" and "Mechanical Backlash" excitations, they were 98.1% and 99.3%, respectively. This is due to the significant number of signals available for testing (532, 283 and 557) respectively. For "Misalignment", the percentage of correct answers was 69.6%. It was also influenced by the error percentages of 11.2% and 19.2%, due to "Unbalance" and

"Mechanical Backlash", respectively. This is due to the low number of signals available for testing, only 28, and because the excitations are related to fr and its harmonics, which causes "Confusion". In a similar way to the previous reasoning, "Mechanical clearance + Misalignment" presented the percentage of correctness of 46.9%, with an error contribution of 40.4% due to "Unbalance". This is also due to the same reasons as before, with the number of signals available for tests being even smaller, that is, only 18. To increase assertiveness, it is necessary to have a greater number of signals, or to apply the "Data Augmentation" method to increase the number of signals or even extract more discriminative parameters in the model. The results show that the proposed methodology allowed the detection of supervised failures in rotating machines, being a promising tool to be applied in Industry 4.0.

Keywords: Fault Detection, Machine Learning, Random Forest, Artificial Intelligence.

1. Introducción

La necesidad de garantizar mayor confiabilidad operativa de los equipos, reducción de costos y mayor disponibilidad para los procesos productivos, hizo que las industrias buscaran estrategias de mantenimiento que minimicen estos problemas. Por su predictibilidad de fallas, el mantenimiento predictivo se ha convertido en una de las mejores alternativas, ya que, a partir de la información recolectada de los equipos, es posible evaluar qué componente de la máquina está defectuoso, verificar la severidad de la falla, realizando la intervención en el momento adecuado, y es posible planificar y programar todos los recursos necesarios para realizar el mantenimiento. Esta estrategia de mantenimiento consiste en técnicas de monitoreo que ocurren periódicamente, los datos recopilados se almacenan para obtener un historial y comparar con las recopilaciones más recientes, analizando la evolución de los síntomas de la máquina monitoreada. El surgimiento de la Industria 4.0 y sus tecnologías hacen que el mantenimiento predictivo sea más eficiente, asertivo y económico, siendo uno de los pilares de esta nueva Revolución Industrial. [1]

El mantenimiento predictivo consiste en varias técnicas de monitoreo de equipos con el poder de intervenir cuando se detecta un problema. Estas técnicas deben permitir la recopilación de datos con el equipo en funcionamiento o con una mínima interferencia en el proceso de producción y deben garantizar que los datos estén sujetos a análisis de tendencia para que se puedan predecir las fallas y, en consecuencia, evitarlas. [2]

El mantenimiento predictivo utiliza varios métodos para monitorear las condiciones de los motores eléctricos aplicados en plantas industriales. Entre estos métodos, el análisis de vibraciones se destaca como un método ampliamente utilizado debido a la posibilidad de identificar una gran posibilidad de fallas. [3]

El análisis de vibraciones se destaca como un método ampliamente utilizado debido a la posibilidad de identificar una amplia gama de fallas. Este proceso aún depende mucho del análisis y procesamiento manual de

la información recolectada, lo que puede requerir más tiempo para realizar el análisis del estado de una máquina, lo que en consecuencia genera más costos para la empresa. [4]

Una de las técnicas más utilizadas dentro del Mantenimiento Predictivo es la detección de fallas a través del análisis de vibraciones, la cual es considerada la Técnica Predictiva más importante para monitorear el estado de las máquinas rotativas en la industria, de todos los parámetros que se pueden medir de manera no invasiva, la cual contiene más la información es la firma de la vibración. Más recientemente, muchas investigaciones han sido motivadas por las aplicaciones de la vibración en el área de la ingeniería, como el diseño de máquinas, cimientos, estructuras, motores, turbinas y sistemas de control. La mayoría de los motores de accionamiento tienen problemas de vibración debido al desequilibrio inherente a los motores. El desequilibrio puede ocurrir debido a fallas en el diseño o mantenimiento deficiente. [5]

La aplicación de técnicas de Inteligencia Artificial en la detección de averías permite realizar un diagnóstico online por parte de un sistema, con una mínima interacción con el usuario y, en muchos casos, pudiendo diagnosticar averías sin la ayuda de especialistas de mantenimiento. El aprendizaje automático es una de las áreas de la informática de más rápido crecimiento con aplicaciones de gran alcance. Se refiere a la detección automatizada de patrones significativos en los datos. [6]

El monitoreo de máquinas rotativas es una tarea esencial en los procesos de producción actuales. Actualmente, varios estudios basados en aprendizaje automático han logrado excelentes resultados en la detección y diagnóstico de fallas. [7]

En este trabajo abordarse la técnica de "Machine Learning" (Aprendizaje Automático), utilizando el clasificador "Random Forest" (Bosque Aleatorio), el cual consiste en una gran cantidad de árboles de decisión individuales que operan como un conjunto. Cada árbol individual en el bosque aleatorio

proporciona una predicción de clase. La clase con más votos se convierte en la predicción del modelo en estudio para una clasificación. De esta manera es posible evaluar si la máquina tiene fallas.

2. Diagnóstico de Fallas

En ese trabajo serán evaluadas las fallas de origen mecánico presentes en un motor de inducción trifásico: Holgura Mecánica; Desalineación Mecánica y Desequilibrio.

2.1. Desequilibrio Mecánico

El desequilibrio mecánico es, con mucho, la mayor causa de vibraciones en las máquinas rotativas. Este tipo de problema que ocurre en las máquinas rotativas está definido por la "ISO - Organización Internacional de Normalización". La vibración resultante es predominantemente radial, fuerte componente a la vez de la frecuencia rotacional ($1 \times f_r$). [8]

Este fenómeno se genera en la mayoría de los casos en el proceso de fabricación debido a asimetrías, desviaciones dimensionales y desviaciones de forma. Cualquiera de estos errores o una combinación de ellos destruirá la condición de perfecta distribución de masa alrededor del eje de rotación del rotor, generando desequilibrio mecánico, que puede ser estático o dinámico.

2.2. Desalineación

La desalineación es un problema casi tan común como el "desequilibrio mecánico". En los conjuntos mecánicos, por lo general, existen varios ejes, cojinetes y acoplamientos con diferentes características dinámicas. Cuando el conjunto opera, aparecen fuerzas de interacción que acaban provocando vibraciones. La desalineación es la desviación de la línea central durante la operación. La desalineación puede ser paralela, angular o una combinación de ambas. [8]

2.3. Holguras Mecánica

La holgura mecánica se caracteriza por la presencia de múltiples armónicos de la frecuencia de rotación ($1 \times f_r$) y genera vibraciones en las máquinas rotativas debido a tornillos flojos, holguras excesivas en los rodamientos, entre otros.

Este tipo de vibración no ocurre sin otras fuerzas que excitan el sistema, como desequilibrio mecánico, desalineación, etc. Cuando existe un juego excesivo, incluso en pequeños desequilibrios mecánicos y desalineaciones, aparecen altos niveles de vibración, es decir, la holgura mecánica amplifica las vibraciones. El plano dominante es el plano radial y la relación de fase es variable. Los armónicos y subarmónicos

superiores de $1 \times f_r$ ($0,5 \times f_r$; $1,5 \times f_r$; $2,5 \times f_r, \dots$) a menudo pueden estar presentes debido a la no linealidad provocada por el truncamiento.

3. "Machine Learning"

El aprendizaje automático es una de las ramas de la inteligencia artificial, con este se logra crear sistemas que pueden aprender de una forma automática e independiente gracias a bases de cualquier tipo de dato previo de una clase en estudio. Es capaz de identificar patrones muy complejos que sin ayuda de un ordenador llevaría bastante tiempo para una persona, los patrones pueden verse en una cantidad muy grande de datos de distintos orígenes, también es posible predecir comportamientos gracias a algoritmos que con el tiempo son capaces de mejorarse de manera independiente. [9]

"Machine Learning" es cuando la máquina utiliza algoritmos para recopilar datos, aprender de parte de ellos y realizar algún análisis o determinación específica que se le ha dado. Por lo tanto, es un campo de estudio que enseña a las computadoras una habilidad que es natural para humanos y animales, a saber, aprender a través de la experiencia sin tener que ser programado explícitamente. Los algoritmos de aprendizaje automático utilizan métodos computacionales para recopilar información directamente de los datos sin depender de una ecuación predeterminada como modelo [6].

En el aprendizaje automático hay dos tipos de técnicas para la aplicación de "Machine Learning" - "Supervised Learning" (Aprendizaje Supervisado), que crea un modelo sobre datos de entrada y salida conocidos para que pueda predecir salidas futuras y "Unsupervised Learning" (Aprendizaje No Supervisado), que encuentra patrones ocultos o estructuras intrínsecas en los datos de entrada [6].

3.1. "Unsupervised Machine Learning"

En el "Unsupervised Machine Learning" (Aprendizaje Automático No Supervisado) existen patrones ocultos o estructuras intrínsecas en los datos. Se utiliza para derivar inferencias de conjuntos de datos que consisten en datos de entrada sin respuestas etiquetadas. En este tipo de aprendizaje no se insertan características / parámetros para encontrar una respuesta. El algoritmo tiene el objetivo de agrupar en "clusters". En este caso el modelo no tiene una respuesta correcta para clasificar, debiendo averiguarlo por sí mismo. [10]

3.2. "Supervised Machine Learning"

En el "Supervised Machine Learning" (Aprendizaje de Máquina Supervisionada), o algoritmo tem un conjunto de datos de entrada e de saída e é treinada

para reconhecer esses padrões. Assim valida-se um modelo que gera previsões sensata para novas entradas de dados. [6]

Básicamente, en " Supervised Machine Learning " los datos consisten en un conjunto de n características, $X = \{X_1, \dots, X_n\}$ donde cada característica X_j se compone de su propio conjunto de valores posibles x_i , con m vectores de características, $x_i = (x_{i1}, \dots, x_{im})$ y una respuesta, $y = (y_1, \dots, y_m)$, que contiene un valor para cada vector, Figura 1.

	X_1	X_2	\dots	X_a	Y
z_1	x_{11}	x_{12}	\dots	x_{1a}	y_1
z_2	x_{21}	x_{22}	\dots	x_{2a}	y_2
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
z_n	x_{n1}	x_{n2}	\dots	x_{na}	y_n

Figura 1. "Supervised Machine Learning". Fuente: [10].

El "Supervised Machine Learning" (Aprendizaje Automático Supervisado) se puede realizar por clasificación o regresión. En este trabajo se abordará el "Classification Method" (Método de Clasificación). "Data Science" (Ciencia de los Datos) proporciona una multitud de algoritmos de clasificación, tales como: "Naive Bayes" (Bayes Ingenuos); "Support Vector Machine" (Máquinas de Vectores Soporte); "Decision Trees" (Árboles de Decisión) y "Random Forest".

3.2.1. "Naive Bayes"

Os "Naive Bayes" son un conjunto de algoritmos de aprendidos supervisados basados en el Teorema de Bayes. Esses algoritmos partem do pressuposto de que existe una independencia condicional entre todos los atributos (ou características) dos ejemplos, dado o contexto da classe, portanto considerados os mais simples dos modelos probabilísticos bayesianos. O que diferencia os tipos de algoritmos "Naive Bayes" é o modo como a função de probabilidade condicional é estimada, isto é, qual a distribución utilizado. [6]

3.2.2. "Support Vector Machine"

La "Support Vector Machine" (Máquina de Vectores de Soporte) fue creada en los años 90 por Cortes y Vapnik y, desde entonces, ha recibido mucha atención por parte de la comunidad de "Machine Learning" debido a su gran poder de generalización en varios dominios y robustez. a datos de alta dimensión. La idea básica es encontrar un hiperplano óptimo que sea capaz de separar los datos de las clases involucradas con un cierto margen máximo. Cuando el problema en cuestión es no lineal y, en consecuencia, no es posible establecer un hiperplano óptimo entre las clases en el propio espacio de entrada. La "Máquina de Vectores de Soporte" utiliza "Funciones del Núcleo" para mapear

los datos del espacio de entrada a un espacio de alta dimensión donde las clases se vuelven linealmente separables. [6]

3.2.3. "Decision Trees"

Una idea por detrás de "Decision Trees" (Árvore de Decisão) es dividir una decisión compleja en una serie de decisiones más simples, con la esperanza de que la solución final obtenida a través de esta estrategia se asemeje a la solución esperada del problema propuesto. Los datos se dividen en regiones más pequeñas según las clases, buscando también minimizar la entropía, es decir, la aleatoriedad de las decisiones tomadas. Un "árbol de decisiones" tiene tres tipos de nodos, como se muestra a continuación. [6]

- Nodo Raíz: El nodo inicial del árbol, donde comienza el proceso de división.

- Nodos de Decisión: Los nodos que contienen cada una de las decisiones tomadas por el algoritmo definen, mediante comparaciones, el camino a seguir por el árbol.

- Nodos Hoja: Los nodos presentes al final del "Árbol de Decisión" representan el resultado final de una combinación de decisiones o eventos.

Inicialmente, el algoritmo busca una característica inicial para iniciar las divisiones y ser el nodo raíz, que en este caso es el que puede dividir el conjunto de datos de la mejor manera posible para que las clases queden bien separadas, pudiendo incluso dividirse en más de dos regiones. Después de dividir, los nodos de prueba se agregan al árbol y se seleccionan nuevas funciones para dividir en el siguiente paso.

3.2.4. "Random Forest"

Los "Random Forest" (Bosques Aleatorios) consisten en una gran cantidad de árboles de decisión individuales que funcionan como un conjunto. Cada árbol individual en el "bosque aleatorio" ofrece una predicción de clase y la clase con más votos se convierte en la predicción del modelo. El concepto fundamental detrás del "bosque aleatorio" es simple pero significativo. En la ciencia de datos, la razón por la que el modelo de "bosque aleatorio" funciona tan bien es la gran cantidad de modelos relativamente no correlacionados ("árboles de decisión") que operan como un comité y superan a cualquiera de los modelos constituyentes individuales. [6]

3.3. Banco de Experimentos

La Figura 2 muestra una descripción general del "Banco Experimental" utilizado para ensamblar la base de datos.

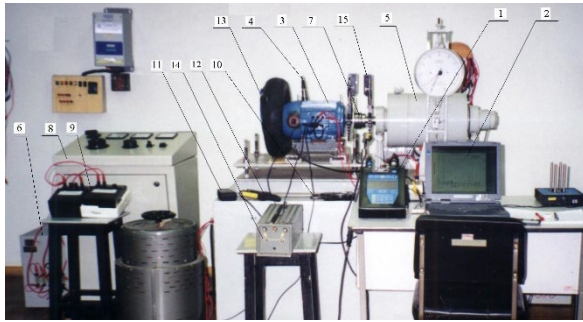


Figura 2. Figuras o tablas a dos columnas insertadas en cuadro de texto. Fuente: [8].

El banco experimental fue instalado en el Laboratorio de Conservación de Energía de la UFSJ - Universidad Federal de São João del Rei. Las señales fueron recolectadas a través del colector y analizador de señales UltraSpec 8000 (1), que digitaliza y memoriza la información contenida en el mismo. Esta información puede ser analizada directamente en el propio equipo o transferida a la computadora (2) a través del software de soporte UltraManager. En esta opción se tiene un ambiente de trabajo más cómodo, permitiendo análisis detallados, generación de informes técnicos y montaje de bases de datos.

Los defectos fueron introducidos en un motor de inducción trifásico (3), WEG (FH 88747), rotor jaula, 5 hp, 1730 rpm, 220 V, 60 Hz, 4 polos, categoría N, 44 barras, 36 ranuras, rodamiento SKF 6205- 2Z, ID-1, carcasa 100L, clase de aislamiento B, FS 1,15, Ip/In 7,5, IP 55, 13,8 A. Se utilizó un generador de CC (5) que alimenta un banco de resistencias (6) como sistema de carga. Se aplicó carga al motor ajustando la corriente de campo del generador de CC. El generador está acoplado al motor eléctrico a través de un acoplamiento flexible (7).

También se utilizó el monitoreo en paralelo utilizando un voltímetro de precisión Engro (9) modelo 600, pinzas digitales Dower modelo CM-600 (12) y un tacómetro Ophtho Tako (4). El propósito de este monitoreo es garantizar el funcionamiento del motor bajo prueba con carga nominal, simulando la condición real de operación. La estructura flexible del banco de experimentación y el motor trabajando a carga nominal permitieron que los defectos introducidos se hicieran más notorios en los espectros. A través de las lecturas del voltímetro (11) se tiene información de los niveles de tensión en las tres fases de alimentación del motor. En el panel de control del generador de CC (8) también es posible monitorear la corriente del inducido, el voltaje del inducido y la velocidad.

3.4. Banco de Datos

Las señales de vibración se recopilaron utilizando el acelerómetro A0720GP, SN6714, sensibilidad de 0,1000 mV/m/s². El analizador se configuró para

ventana de Hanning, 3200 líneas, 10 promedios, rango de frecuencia de 0 a 400 Hz y amplitudes medidas en velocidad (mm/s). Para observar los defectos que se manifiestan a altas frecuencias, se recolectaron algunos espectros en el rango de frecuencia de 0 a 1200 Hz. El acelerómetro se colocó alternativamente en las direcciones vertical, horizontal y axial, lado ventilador y lado acoplamiento motor.

Se recolectaron aleatoriamente 3916 espectros en el módulo UltraSpec Analyzer (3492 - análisis de vibración, 212 - análisis de corriente y 212 - análisis de flujo magnético), en una serie de 50 pruebas para cada excitación y repetidas bajo las mismas condiciones. En este trabajo solo se utilizaron los espectros de análisis y vibración.

3.5. Parámetros

Los parámetros fueron seleccionados con base en el conocimiento disponible sobre los tipos de fallas en estudio, como se muestra en las Figuras 3 a 7. Se puede observar que existen diferencias entre los parámetros extraídos, sin embargo, no son totalmente discriminativos.

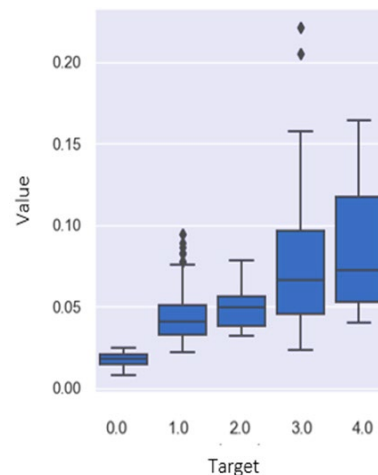


Figura 3. RMS. Fuente: Elaboración propia.

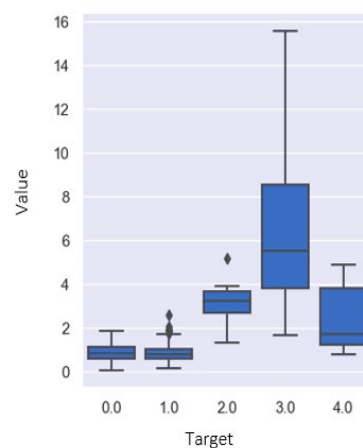


Figura 4. Energía 1 x f. Fuente: Elaboración propia.

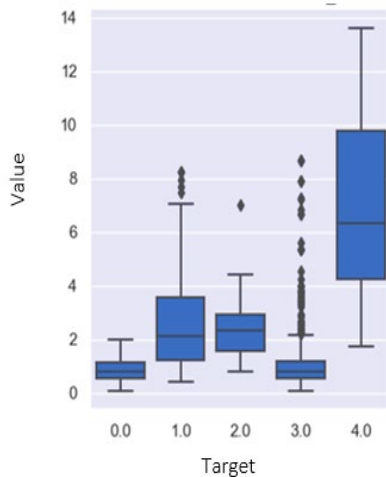


Figura 5. Energía 2 x f_r . Fuente: Elaboración propia.

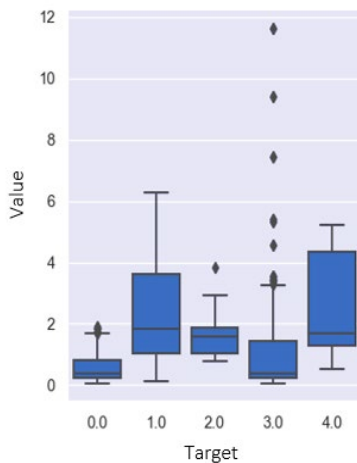


Figura 6. Energía 3 x f_r . Fuente: Elaboración propia.

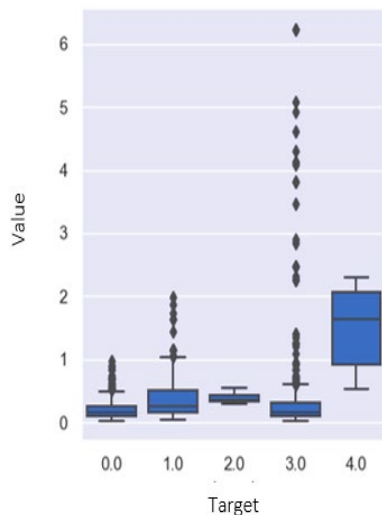


Figura 7. Energía 4 x f_r . Fuente: Elaboración propia.

3.6. Señales

Para el modelo presentado, se utilizarán las señales que se presentan a continuación.

- Sin defecto (Etiqueta = 0), 532 señales.
- Desequilibrio (Etiqueta = 1), 283 señales.
- Desalineación (Etiqueta = 2), 28 señales.
- Holgura Mecánica (Etiqueta = 3), 557 señales.
- Holgura Mecánica e Desalineación (Etiqueta = 4), 18 señales.

La velocidad de rotación se mantuvo constante a 1717,5 rpm. Las observaciones se etiquetaron de acuerdo con la condición impecable y las otras cuatro excitaciones antes mencionadas se introdujeron en el "Banco de Experimentos". Posteriormente, se realizó el análisis del espectro de vibración de cada uno de ellos. Cada archivo consta de 3200 puntos con $d_f = 0,125$ Hz.

3.7. Machine Learning

En este trabajo se utilizó el modelo "Random Forest", Figura 8. El modelo del algoritmo fue creado de acuerdo con los parámetros que se presentan a continuación.

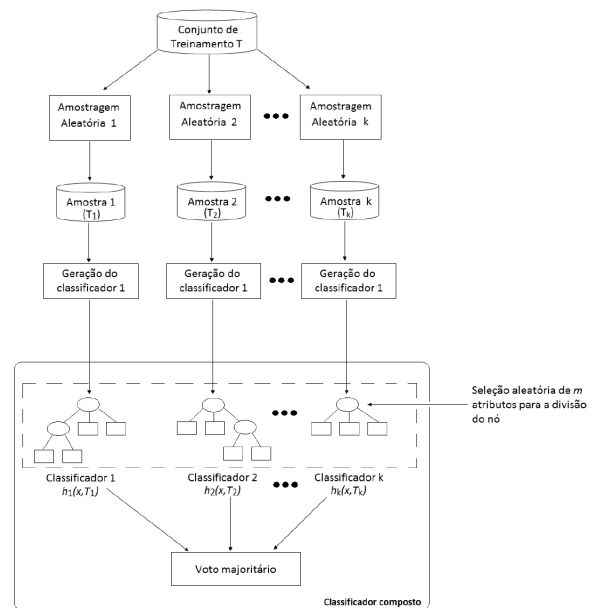


Figura 8. "Rondon Forest". Fuente: [10].

- Normalizado basado en el grupo de entrenamiento. Usando la función "StandardScaler".
- Se dividieron en un 70% para entrenamiento y un 30% para pruebas.

- Los análisis se realizaron 1000 veces, dividiendo aleatoriamente los grupos de entrenamiento y prueba para verificar la robustez del modelo y los resultados obtenidos. Además, con cada nueva división, el algoritmo también se iniciaba aleatoriamente.

- El algoritmo utilizado es "Random Forest" con aproximadamente 200 iteraciones para la condición sin falla y las otras cuatro excitaciones. Los

hiperparámetros del modelo se ajustaron con un breve análisis de las señales del grupo de entrenamiento.

- Los resultados se presentan en cuanto a la matriz de confusión por el desequilibrio de los datos, además de la precisión y desviación estándar para el análisis global, aún sabiendo que para casos desequilibrados no es la mejor métrica, pero como tiene la matriz de confusión con la clasificación por tipo de falla que podemos considerar.

As excitações com menor quantidade de sinais tendem a ter erros maiores. Isso deve pelo fato de o algoritmo não conseguir aprender tão bem quanto aqueles que possuem muitos sinais, conforme será mostrado na "Matriz de Confusão" apresentada no item 3.8.

3.8. Matriz de Confusão

La "Matriz de Confusión", Figura 9, se calcula mediante la función de búsqueda "Rango". Muestra la distribución de registros en términos de sus clases actuales y sus clases pronosticadas. Esto indica la calidad del modelo actual.

Matriz de Cofusão Média Porcentagem:				
Normal	99.9	0.1	0.0	0.0
Desbal	0.5	98.1	0.6	0.8
Desali	0.11.2	69.6	19.2	0.1
Folga	0.2	0.2	0.3	99.3
Fol+des	0.4	0.4	5.5	7.3
N / Des / Desal / Folg / Folg+Des				
Accuracy Média: 98.06 % Desvio: 0.65 %				

Figura 9. " Matriz de Confusión " del Modelo "Rondon Forest". Fuente: Elaboración propia.

3.9. Resultados

Los resultados obtenidos con la metodología propuesta para la clasificación supervisada "Supervised Machine Learning" (Aprendizaje Automático Supervisado) se presentan en la Figura 10, de acuerdo con las "Etiquetas" presentadas en el ítem 3.6 "Señales".

La condición "Sin Defecto", representada en la Figura 10 en verde ("Normal" = 99 %) y las excitaciones representadas en naranja ("Desequilibrio" = 98,1 %) y lila ("Holgura Mecánica" = 99,3 %) mostraron excelentes resultados. Esto se debe al importante número de señales disponibles para la prueba, es decir, 532, 283 y 557, respectivamente.

Con respecto a la excitación "Desalineación", dos comentarios son necesarios para una mejor comprensión de los resultados. El primero está relacionado con el porcentaje de aciertos, representado en la Figura 10 por el color azul, que es del 69,6%. Esto se debe a que el número de señales es bajo, es decir,

solo 28. El segundo está relacionado con el porcentaje de error, representado en rojo, con un 11,2 % por "Desequilibrio" y un 19,2 % por "Holgura Mecánica" . Además de tener solo 28 señales para probar, estas excitaciones están relacionadas con f_r y sus armónicos, lo que provoca "Confusión"

Similar al razonamiento anterior es el resultado relativo al "Holgura Mecánica + Desalineación" que presentó el porcentaje de acierto del 46,9%. El error del 40,4% se debe a "Desequilibrio" (color negro). Esto se debe a que tiene solo 18 señales para tets y estas excitaciones están relacionadas con frecuencia de rotación (f_r) y sus armónicos lo que provoca "Confusión".

Para aumentar la asertividad, es necesario tener un mayor número de señales, o aplicar el método "Data Augmentation" (Data Augmentation), para aumentar el número de señales o incluso extraer más parámetros discriminativos en el modelo. El "aumento de datos" no es el objetivo de este trabajo y será abordado en trabajos futuros.

3.10. Conclusiones

En este trabajo se presentó un nuevo enfoque para la detección y diagnóstico de fallas en máquinas rotativas, a través del "Supervised Machine Learning", (Aprendizaje Automático Supervisado).

Las características de vibración en el "Dominio del Tiempo" y "Espectro de Frecuencia" fueron extraídas con base en el conocimiento humano que integra este Equipo de Gep_LASID (Grupo de Estudio e Investigación del Laboratorio de Sistemas Dinámicos), del DEMEP (Departamento de Ingeniería Mecánica y de Producción), de la UFSJ (Universidad Federal de São João del-Rei). En el diagnóstico se verificó la presencia de la condición sin falla y las otras cuatro excitaciones de forma "Supervisada" en base al algoritmo "Random Forest".

Los resultados muestran que la metodología propuesta permite la detección de fallas supervisadas en máquinas rotativas.

Ya se está construyendo una nueva "Base de Datos" para la aplicación de "Aumento de Datos", además de otras técnicas que el Equipo Gep_LASID está investigando. El objetivo es seguir contribuyendo a la detección de defectos incipientes en máquinas rotativas a través de Técnicas Predictivas e Inteligencia Artificial con un enfoque de Industria 4.0.

4. Referencias

[1] C. A. Pires, R. H. Okada. "Manutenção Preditiva: Estratégia de produção e redução de custos". Interface Tecnológica - v. 17 n. 1, 2020 [PDF]. Disponible:

<https://revista.fatectq.edu.br/interfacetecnologica/article/view/781/502>.

[2] V. C. Souza, C. S. Marchi, N. V. Bueno, T. S. Faustino, T. A. Barreiro. "Utilização das tecnologias da indústria 4.0 na manutenção preditiva através do monitoramento de equipamentos e instalações". *Brazilian Journal of Development*, Curitiba, v.8, n.1, p. 7063-7083 jan. 2022.

[3] A. S. Araújo. "Análise, avaliação e validação do uso de técnicas de inteligência artificial para diagnóstico de falhas em motores de indução. Dissertação (Mestrado). 88 f. 2022 [PDF]. Disponível: <http://repositoriosenaiba.fieb.org.br/handle/fieb/1231>.

[4] Holanda, S. M. S. "Aplicação da Manutenção Preditiva por análise de vibrações em equipamentos de trens urbanos com plano de manutenção proposto". Recife, 2016.

[5] L. C. Brito. "An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery. *Mechanical Systems and Signal Processing*". [PDF]. Disponível: www.elsevier.com/locate/ymssp, consulta realizada no dia 24/04/2022.

[6] Osisanwo F. Y., Akinsola J. E .T., Awodele O., Hinmikaiye J. O., Olakanmi, O., Akinjobi, J. "Supervised machine learning algorithms: classification and comparison". *International Journal of Computer Trends and Technology (IJCTT)*, v. 48, n. 3, p. 128-138, 2017.

[7] Matos, G, R. "Machine Learning aplicado à Gestão de Activos Físicos Industriais". Instituto Superior de Engenharia de Lisboa, Departamental de Engenharia Mecânica. Dissertação (Mestrado), 2021, 89 P.

[8] J. N. Brito. "Desenvolvimento de um sistema inteligente híbrido para diagnóstico de falhas em motores de indução trifásicos". 2002. 214p. Tese (Doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica, Campinas, SP.

[9] M. G. Gramajo, L. Ballejos and M. Ale, "Software Requirements Engineering through Machine Learning Techniques: A Literature Review," 2018 IEEE Biennial Congress of Argentina (ARGENCON), 2018, pp. 1-7, doi: 10.1109/ARGENCON.2018.8646010.

[10] T. M. Oshiro. "Uma abordagem para a construção de uma única árvore a partir de um Random Forest para classificação de bases de expressão gênica". Universidade de São Paulo. Programa Inter unidades em Bioinformática. Dissertação (Mestrado). 101 p. 2013.

[11] L. Silva, K. Utimura, M. C. Silva, S. Prado. "Study on Machine Learning Techniques for Botnet Detection". *IEEE Latin America Transactions*. Vol. 18, No 5, MAY 2020, 8p.