

Incremento del tamaño de los datos para la detección de fallos en maquinaria rotativa

Data size increment for fault detection on rotating machinery

Adriana Guamán Buestán¹.,Diego Cabrera Mendieta²., René V. Sanchez Loja³., Mariela Cerrada Lozada⁴., Juan Manuel Cevallos Ampuero ⁵

¹ MBA.Docente investigadora. Universidad Politécnica Salesiana, grupo GIDTEC. aguaman@ups.edu.ec

² Ph.D. Docente investigador. Universidad Politécnica Salesiana, grupo GIDTEC. dcabrera@ups.edu.ec

³ Ph.D. Docente investigador. Universidad Politécnica Salesiana, grupo GIDTEC. rsanchez@ups.edu.ec

⁴ Ph.D.Docente investigadora. Universidad Politécnica Salesiana, grupo GIDTEC. mcerrada@ups.edu.ec

⁵ Ph.D.Director Instituto de Posgrados. Universidad Nacional Mayor de San Marcos. juanmanuelcevallos@gmail.com

Resumen

En los últimos años se ha incrementado el uso de técnicas de modelamiento basado en datos para el diagnóstico de fallos en maquinaria rotativa. Estas técnicas requieren de grandes cantidades de datos que no siempre se pueden obtener pues generan altos costos y tiempo excesivo, que son difíciles de solventar desde el punto de vista económico y técnico. El presente trabajo se enfoca en el pre-procesamiento de las señales de vibración y propone un método para incrementar el número de series temporales informativas de una máquina rotativa sin el incremento del tiempo y costos en la etapa de adquisición de las señales. Como resultado se ha obtenido una ampliación de 315 señales en la fase de adquisición de datos a 429000 luego de la aplicación del método; cantidad adecuada para la construcción de modelos basados en datos, incluso de deep learning para la detección de fallos en maquinaria rotativa.

Palabras clave: Adquisición de datos, pre-procesamiento, rodamientos, señales.

Abstract

In recent years, the use of data-based modeling techniques for the diagnosis of rotating machinery failures has increased. These techniques require large amounts of data that can't always be obtained because they generate high costs and excessive time, which are difficult to solve from the economic and technical point of view. The present work focuses on the pre-processing of vibration signals and proposes a method to increase the number of informative time series of a rotary machine without the increase of time and costs in the stage of acquisition of signals. As a result, an expansion of 315 signals was obtained in the data acquisition phase to 429000 after the application of the method; adequate amount for the construction of models based on data, including deep learning for the detection of failures in rotating machinery.

Keywords: Data acquisition, pre-processing, bearings, signals.

1. INTRODUCCIÓN

La evolución del mantenimiento desde varias décadas, ha impulsado el desarrollo del

mantenimiento predictivo (Ahmad y Kamaruddin, 2012). Al respecto se ha dado una gran cantidad de aportes que le han posicionado como un factor clave dentro de los procesos productivos y ser visto como una oportunidad de mejora por la reducción de costos por paradas no programadas, seguridad para el operario y mejora en la calidad de los productos (Jardine, Lin, y Banjevic, 2006). Los esfuerzos en el ámbito investigativo para la detección de fallos, como los realizados por Yin, Ding, Xie, Luo, (2014), Goyal y Pabla, (2016) y Cabrera, Sancho, Cerrada, Sánchez, y Tobar, (2018) han dado como resultado el avance de diversas estrategias para el monitoreo de la condición, con el uso de datos de parámetros como la vibración, emisión acústica, entre otras.

Para estas estrategias es necesario contar con las señales que definan la condición de la máquina. Su adquisición supone ser la etapa que mayor tiempo y recursos consume, pues las señales adquiridas deben ser lo más realistas posibles en condiciones específicas establecidas y que contengan las características deseadas para el modelamiento (Virtanen, Plumbley, y Ellis, 2017). En muchos casos esto resulta poco viable por los altos costos y tiempo necesario. El motivo que lo genera es la diversidad de fallos que se pueden originar, considerando que se presentan de forma individual o secuencial.

El proceso de adquisición de señales involucra la recopilación y almacenamiento de información proveniente de elementos específicos que son de interés del investigador como la elección del tipo de sensor, su ubicación, cantidad, el hardware de adquisición y almacenamiento, así como el procedimiento para su adquisición (Fugate, Sohn, y Farrar, 2001). Al necesitar una cantidad robusta de señales este procedimiento se vuelve largo y costoso pudiendo ser en muchos casos inviable.

También es importante puntualizar que para el modelamiento basado en datos es necesario realizar la clasificación de las características que identifican el estado de la máquina. Para el efecto, existen métodos que se ejecutan de forma manual y automática (Kumar, Shankar, y Thakur, 2017). Para los métodos basados en aprendizaje automático, se requiere una cantidad suficiente de datos que proporcionen diversidad de las diferentes condiciones de la máquina y que permitan que el modelo tenga capacidad de generalización.

Por lo anterior, este documento propone una metodología para el pre-procesamiento de las señales de vibración y la extracción de características como insumo para la detección de fallos de maquinaria rotativa. De esta manera, se propone contribuir con datos suficientes para la generación de un modelo y que sirva como base para futuras investigaciones en el campo del mantenimiento basado en la condición.

2. METODOLOGÍA

Las señales adquiridas en una fase previa desde la maquinaria, requieren ser pre-procesadas, pues la intención es incrementar su cantidad y eliminar perturbaciones o valores atípicos que pudieran interferir en el modelamiento. En la Figura 6 se esquematiza los pasos que se explican a continuación.

2.1 Ingreso de las señales

Como ya se mencionó, en un proceso previo se toman las señales de vibración desde la maquinaria rotativa y éstas pueden ingresar directamente desde los dispositivos de adquisición para el pre-procesamiento o también se pueden almacenar en una base de datos.

2.2 División del conjunto de señales

43

El siguiente paso es la división del conjunto de señales en: señales de entrenamiento, señales de validación y señales de prueba. Esta acción es necesaria, pues en la creación del modelo no se pueden utilizar todos los datos. Es necesario contar con un subconjunto que se pueda aplicar al momento de la generación del modelo, y otros dos para su validación. En esta fase también se extrae la clase de cada señal de los tres subconjuntos. Es importante puntualizar que la clase hace referencia al nivel de severidad del fallo, dato necesario para el proceso de clasificación al momento de crear un modelo.

2.3 Ventaneo de la Señal

Las series de tiempo de cada subconjunto son extensas como la que se muestra en la Figura 1, por lo tanto es necesario seccionar en intervalos más pequeños a través de un ventaneo.

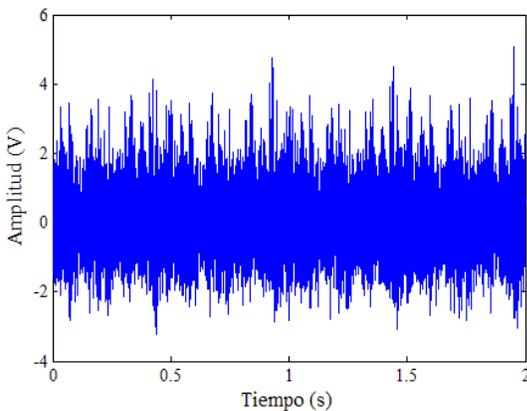


Figura 1. Señal de vibración

Para este procedimiento se requiere establecer parámetros de ajuste como: el incremento (ΔT), ventana (w) y ancho (ts). Cada parámetro cumple una función específica que en conjunto permitirá seccionar la señal en cuantas partes resulte por los valores establecidos. De este modo, el incremento determina cuánto se mueve la ventana a lo largo de la señal, la ventana especifica el ancho de seccionamiento primario de la señal y el ancho

indica en cuántas partes se efectúa el seccionamiento secundario; dicho de otro modo, este último parámetro denota el tamaño de las subseñales resultantes.

Para comprender mejor lo descrito, se ilustra en la Figura 2 una porción de la señal. El recuadro en naranja indica el tamaño del seccionamiento primario (w) y el incremento detalla cuánto se va a mover la ventana entre cada seccionamiento. Para esta señal se tiene un incremento (ΔT) igual a cinco que establece que la ventana se desplazará con sobrelapamiento por cinco ocasiones, pues el ancho del intervalo no es igual al ancho de la ventana. El resultado de esta operación es un seccionamiento primario de la señal en cinco partes (Figura 3).

La señal resultante del seccionamiento aún es extensa, por tanto es necesario realizar un seccionamiento secundario. Para el efecto se toma cada ventada del seccionamiento primario y dentro de éstas, se genera una subdivisión con un ancho definido por ts (Figura 4).

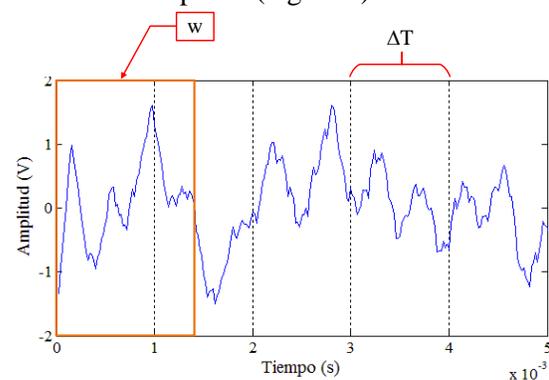


Figura 2. Parámetros de ajuste primario

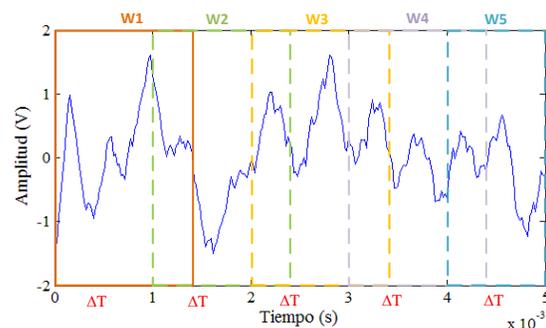


Figura 3. Ventaneo primario

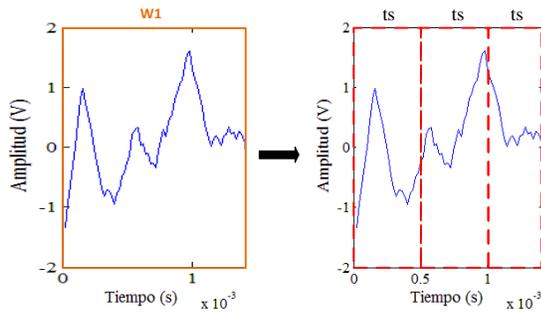


Figura 4. Ventaneo secundario

El resultado de esta operación es el incremento de señales por cada ventana de seccionamiento. Así para el ejemplo que se ha ilustrado, inicialmente se hizo un primer seccionamiento en cinco partes, posteriormente de cada ventana se tomó tres subseñales obteniendo al final un total de 15. De esta forma, el procedimiento aplicado ha permitido obtener un conjunto de subseñales a partir de una señal, como se esquematiza en la Figura 5. Es oportuno recordar que el incremento se efectúa en cada señal de los tres subconjuntos.

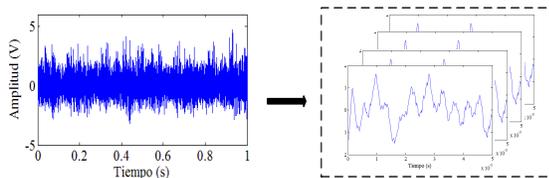


Figura 5. Resultado del incremento de señales

Luego del incremento de señales, se procede a calcular las características estadísticas que según la bibliografía consultada (Sanchez, 2017) son atributos útiles para la generación de un modelo, ésta son: media, desviación estándar, curtosis, raíz cuadrada media, valor pico positivo, factor de cresta, r media, factor de forma, factor de impulso, varianza y valor mínimo.

2.4 Estandarización de la señal

Para terminar el proceso se estandariza la señal con la finalidad de eliminar valores que no están

dentro del rango normal y que pudieran alterar los resultados en el proceso de modelado (principalmente para redes neuronales). Para el efecto se ajusta la señal a una distribución normal a través de la siguiente relación.

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Donde:

z, señal normalizada

x, señal a normalizar

μ , media

σ , varianza

Es importante puntualizar que este procedimiento se aplica únicamente al conjunto de señales de entrenamiento, pues si se toma las de validación y prueba se estaría sesgando el resultado del modelamiento. Esto se da porque en la creación del modelo se utiliza únicamente el conjunto de entrenamiento; no obstante, los valores de la media y de la desviación estándar se aplican también a los otros dos subconjuntos.

Al finalizar este procedimiento, se obtienen tres subconjuntos de entrenamiento, validación y prueba con la cantidad adecuada de señales estandarizadas que están listas para realizar un posterior modelamiento. En la Figura 6 se resume lo pasos que se han descrito para el incremento de las señales de vibración.

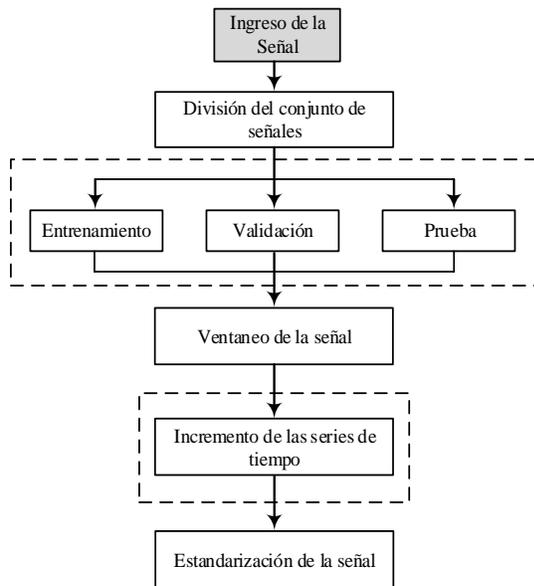
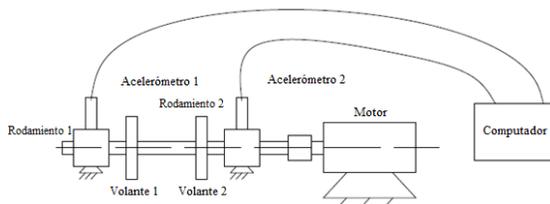


Figura 6. Procedimiento para el incremento de las señales de vibración. Los rectángulos interpunteados representan resultados y el rectángulo sombreado indica que los datos provienen de otro procedimiento.

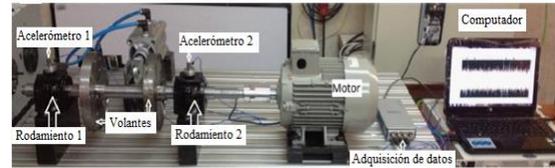
3. EXPERIMENTO

3.1 Base de datos e ingreso de la señal

Para la experimentación se utilizó la base de datos del laboratorio de vibraciones de la Universidad Politécnica Salesiana sede Cuenca, Ecuador. Los datos adquiridos previamente provienen del banco de pruebas mostrado en la Figura 7; siendo importante señalar que la metodología considera las mediciones realizadas en los dos rodamientos.



a) Configuración esquemática



b) Configuración real

Figura 7. Banco de vibraciones

Adicionalmente, la condición de la máquina está referenciada por siete niveles de severidad que se estructuran como se muestra a continuación.

Tabla 1. Niveles de severidad

Nivel de Severidad	Elementos evaluados	
	Rodamiento 1	Rodamiento 2
P1	Normal	Normal
P2	Fallo en pista interna	Normal
P3	Fallo en pista externa	Normal
P4	Fallo en elemento rodante	Normal
P5	Fallo en pista interna	Fallo en pista externa
P6	Fallo en pista interna	Fallo en elemento rodante
P7	Fallo en elemento rodante	Fallo en elemento rodante

De acuerdo a la tabla 1, el nivel P1 indica que los dos rodamientos están en condición normal, no así el nivel P7 donde se especifica que el rodamiento 1 tiene un fallo en su elemento rodante en combinación con un fallo en el elemento rodante del rodamiento 2. Esta configuración determina la condición de la maquinaria.

Para aplicar la metodología propuesta también es necesario conocer las condiciones establecidas en la adquisición de las señales de vibración. Para el efecto se han realizado cinco repeticiones (R1, R2, R3, R4, R5), con tres frecuencias 8Hz (F1), 10Hz (F2), 15Hz (F3) y con tres tipos de carga

46

(L1, sin carga), (L2, con una carga) y (L3, con dos cargas). Esta configuración experimental ha orientado a la codificación para cada señal en la base de datos. En consecuencia cada una está identificada por la repetición, frecuencia, carga y tipo de fallo. Así la señal R1F1L1P1 detalla que corresponde a la primera repetición, con una frecuencia de 8Hz, sin carga, y con nivel de severidad 1, es decir en condición normal. De esta codificación el último dígito es muy importante, pues se utiliza para la extracción de la clase (nivel de severidad). Con la codificación descrita, el conjunto de señales contenidas en la base de datos son pre-procesadas con la metodología propuesta.

3.2 División del conjunto de señales

La división de los datos se estableció de la siguiente forma: datos de entrenamiento con un 70% del total, datos de validación con el 15% y datos de prueba con el 15% restante. Este proceso se realizó de forma estratificada y aleatoria para proporcionar una distribución homogénea de las señales, de tal forma que se obtengan subconjuntos con la cantidad adecuada y garantizando la presencia de señales con todos niveles de severidad en cada segmento.

3.3 Ventaneo de la Señal

Para el seccionamiento de la señal fue necesario establecer los valores de los parámetros de ajuste que fueron los siguientes:

Ancho (ts) = 100

Incremento (ΔT) = 250

Ventana (w) = 500

3.4 Estandarización

Para la estandarización, a cada señal del conjunto de entrenamiento que correspondió al 70% del total de datos, se aplicó la relación indicada en la

sección 2.4, obteniendo una señal ajustada a una distribución normal. Los valores de la media y desviación estándar se hicieron extensivos para los otros dos subconjuntos.

4. RESULTADOS

Del proceso previo en la adquisición de señales se ha obtenido un total de 315 señales en el dominio tiempo (Figura 8). Esta cantidad proviene del número de pruebas efectuadas por cada nivel de severidad (clase). De esta forma al tener siete niveles, por cada uno se han obtenido 45 señales dando como resultado la cantidad arriba indicada. Con este conjunto de datos se aplicó el pre-procesamiento.

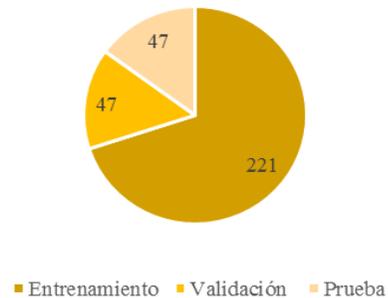


Figura 8. Cantidad inicial de señales por subconjunto

Luego del ventaneo finalmente se obtuvo una cantidad de 429000 señales que están distribuidas de la siguiente forma.

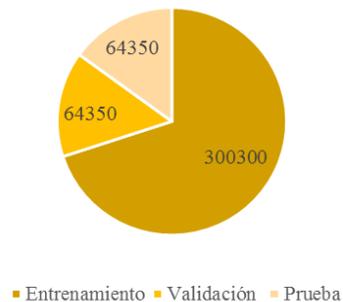


Figura 9. Cantidad de señales por subconjunto después del ventaneo

47

Finalmente, con el proceso de estandarización las señales de entrenamiento fueron ajustadas a una distribución normal standard, donde su desviación estándar es 0 y su varianza 1.

4. CONCLUSIONES

El proceso descrito ha permitido incrementar el tamaño de datos de 315 señales a 429000 en total y contienen iguales características estadísticas que las originales. Esta metodología aporta en la disminución de tiempo y costo al momento de obtener la cantidad de datos adecuada para la generación de un modelo para la detección de fallos en maquinaria rotativa, sin descartar su uso para otro tipo de maquinaria.

La cantidad de subseñales a obtener depende del valor de los parámetros de ajuste y tiene relación directa con el requerimiento que se proponga para el modelado.

Las señales obtenidas en este procedimiento están estandarizadas y no requieren otro paso previo para la modelización.

Referencias Bibliograficas

Ahmad, R., y Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering*, 63(1), 135-149. <https://doi.org/10.1016/j.cie.2012.02.002>

Cabrera, D., Sancho, F., Cerrada, M., Sánchez, R.-V., y Tobar, F. (2018). Echo state network and variational autoencoder for efficient one-class learning on dynamical systems. *Journal of Intelligent & Fuzzy Systems*, 34(6), 3799-3809. <https://doi.org/10.3233/JIFS-169552>

Fugate, M. L., Sohn, H., y Farrar, C. R. (2001). VIBRATION-BASED DAMAGE DETECTION USING STATISTICAL PROCESS CONTROL. *Mechanical Systems and Signal Processing*, 15(4), 707-721. <https://doi.org/10.1006/mssp.2000.1323>

Goyal, D., y Pabla, B. S. (2016). The Vibration Monitoring Methods and Signal Processing Techniques for Structural Health Monitoring: A Review. *Archives of Computational Methods in Engineering*, 23(4), 585-594. <https://doi.org/10.1007/s11831-015-9145-0>

Jardine, A. K. S., Lin, D., y Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7), 1483-1510. <https://doi.org/10.1016/j.ymsp.2005.09.012>

Kumar, A., Shankar, R., y Thakur, L. S. (2017). A big data driven sustainable manufacturing framework for condition-based maintenance prediction. *Journal of Computational Science*. <https://doi.org/10.1016/j.jocs.2017.06.006>

Sanchez, L., Vinicio, R. (2017). Diagnóstico de fallos en cajas de engranajes mediante la aplicación de diferentes técnicas de inteligencia artificial. Recuperado de <http://espacio.uned.es/fez/view/tesisuned:ED-Pg-TecInd-Rvsanchez>

Virtanen, T., Plumbley, M. D., y Ellis, D. (2017). *Computational Analysis of Sound Scenes and Events*. Springer.

Yin, S., Ding, S. X., Xie, X., y Luo, H. (2014). A Review on Basic Data-Driven Approaches for Industrial Process Monitoring. *IEEE Transactions on Industrial Electronics*, 61(11), 6418-6428. <https://doi.org/10.1109/TIE.2014.2301773>

*Para citar este artículo:Guamán Buestán A., Cabrera Mendieta. D., Sanchez Loja R.V., Cerrada Lozada M., Cevallos Ampuero J.M. data size increment for fault detection on rotating machinery . Revista Bistua.2019. 17(2):41-48.

+ Autor para el envío de correspondencia y la solicitud de las separatas : Guamán Buestán A. Docente investigadora. Universidad Politécnica



48

Salesiana, grupo GIDTEC. guaman@ups.edu.ec

Recibido: Septiembre 05 de 2018

Aceptado: Enero 20 de 2019

Finalmente, con el proceso de estandarización las señales de entrenamiento fueron ajustadas a una distribución normal standard, donde su desviación estándar es 0 y su varianza 1.

4. CONCLUSIONES

El proceso descrito ha permitido incrementar el tamaño de datos de 315 señales a 429000 en total y contienen iguales características estadísticas que las originales. Esta metodología aporta en la disminución de tiempo y costo al momento de obtener la cantidad de datos adecuada para la generación de un modelo para la detección de fallos en maquinaria rotativa, sin descartar su uso para otro tipo de maquinaria.

La cantidad de subseñales a obtener depende del valor de los parámetros de ajuste y tiene relación directa con el requerimiento que se proponga para el modelado.

Las señales obtenidas en este procedimiento están estandarizadas y no requieren otro paso previo para la modelización.

Referencias Bibliograficas

Ahmad, R., y Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering*, 63(1), 135-149. <https://doi.org/10.1016/j.cie.2012.02.002>

Cabrera, D., Sancho, F., Cerrada, M., Sánchez, R.-V., y Tobar, F. (2018). Echo state network and variational autoencoder for efficient one-class learning on dynamical systems. *Journal of Intelligent & Fuzzy Systems*, 34(6), 3799-3809. <https://doi.org/10.3233/JIFS-169552>

Fugate, M. L., Sohn, H., y Farrar, C. R. (2001). VIBRATION-BASED DAMAGE DETECTION USING STATISTICAL PROCESS CONTROL. *Mechanical Systems and Signal Processing*, 15(4), 707-721. <https://doi.org/10.1006/mssp.2000.1323>

Goyal, D., y Pabla, B. S. (2016). The Vibration Monitoring Methods and Signal Processing Techniques for Structural Health Monitoring: A Review. *Archives of Computational Methods in Engineering*, 23(4), 585-594. <https://doi.org/10.1007/s11831-015-9145-0>

Jardine, A. K. S., Lin, D., y Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7), 1483-1510. <https://doi.org/10.1016/j.ymssp.2005.09.012>

Ramírez AL, Gil J, Medina MH, Cruz B. (2016). Implementación en entorno Labview de un sistema multifuncional de medidas magnetoópticas y magnetoeléctricas para caracterización de materiales. *BISTUA Revista de la Facultad de Ciencias Básicas*, 14 (2): 116-125. doi: <https://doi.org/10.24054/01204211.v2.n2.2016.2188>

Kumar, A., Shankar, R., y Thakur, L. S. (2017). A big data driven sustainable manufacturing framework for condition-based maintenance prediction. *Journal of Computational Science*. <https://doi.org/10.1016/j.jocs.2017.06.006>

Sanchez, L., Vinicio, R. (2017). Diagnóstico de fallos en cajas de engranajes mediante la aplicación de diferentes técnicas de inteligencia artificial. Recuperado de <http://e-spacio.uned.es/fez/view/tesisuned:ED-Pg-TecInd-Rvsanchez>

Virtanen, T., Plumbley, M. D., y Ellis, D. (2017). *Computational Analysis of Sound Scenes and Events*. Springer.

Yin, S., Ding, S. X., Xie, X., y Luo, H. (2014). A Review on Basic Data-Driven Approaches for Industrial Process Monitoring. *IEEE Transactions on Industrial Electronics*, 61(11), 6418-6428. <https://doi.org/10.1109/TIE.2014.2301773>



*Para citar este artículo:Guamán Buestán A.,
Cabrera Mendieta. D., Sanchez Loja R.V.,
Cerrada Lozada M., Cevallos Ampuero .J.M. data
size increment for fault detection on rotating
machinery. Revista Bistua.2019. 17(2):41-48.

+ Autor para el envío de correspondencia y
la solicitud de las separatas : Guamán Buestán
A. *Docente investigadora. Universidad
Politécnica
Salesiana, grupo GIDTEC. guaman@ups.edu.ec*

Recibido: Septiembre 05 de 2018

Aceptado: Enero 20 de 2019

