

Aplicación de machine learning en el mantenimiento predictivo industrial con herramientas de código abierto

Jorge Ivan Romero Gelvez, Brayan Stiven Rincón Quintero

Universidad de Bogotá Jorge Tadeo lozano, Bogotá, Colombia

jorgei.romerog@utadeo.edu.co

Resumen

Este trabajo muestra la aplicación del problema de mantenimiento predictivo industrial usando un conjunto de datos que se utilizara para la detección de anomalías utilizando código abierto. El objetivo del modelo es detectar anomalías de picos y puntos de cambio en los datos en mantenimiento. El paso inicial es la selección de un conjunto de datos extenso con el cual se estudiar el comportamiento de la variable en el tiempo, identificando si este comportamiento se ve afectado por una tendencia o un componente estacional. Aquí el objetivo es definir el comportamiento esperado o normal.

Luego se debe identificar y estudiar patrones de comportamiento raros e inusuales y validar si pueden afectar el patrón normal de alguna manera. Lograr encontrar los valores atípicos y anomalías puede significar en algún tipo de problema o consecuencia de un problema el cual se debe tratar de manera oportuna. Esto se logra al entrenar los datos para posterior generar un modelo y este debe ser evaluado para obtener la calidad del modelo.

Como punto de comparación se decidió utilizar la optimización de enjambre de partículas con la cual encontramos un error cuadrático medio más bajo que el anteriormente obtenido con fb-prophet; por lo cual se concluyó que para usos futuros este algoritmo ofrece mas confiabilidad.

Palabras claves: aprendizaje automático, detección de anomalías, código abierto, mantenimiento, predictivo, optimización.

Introducción

El futuro del mantenimiento empresarial es digital. Los avances de la tecnología e internet de las cosas está acoplado muchos de los dispositivos y maquinarias de mayor uso hoy en día gracias al uso de sensores conectados a servidores distantes que almacenan estos datos, lo que permite a técnicos de mantenimiento e interesados tener a disposición datos en tiempo real del estado de la maquinaria industrial; esto facilita dejar atrás conceptos de mantenimiento correctivo y preventivo para dejar en primer lugar el

mantenimiento predictivo por lo cual este trabajo está enfocado en la detección de anomalías con lo cual las empresas generaran programación de mantenimiento para sus industrias, reduciendo la cantidad de sobrecostos por daños y averías.

Una de las preocupaciones en la industria es el aumento de conexión entre máquina y la nube por la reglamentación de la ISO 27001 sistema de gestión de seguridad de la información en la cual se establece parámetros de seguridad que pueden ser vulnerados por medio de múltiples métodos de hacking; teniendo en cuenta que los datos industriales ya sean instrucciones, métodos, datos y secretos industriales son uno de los recursos más vitales para muchas industrias.

Por lo cual las empresas han estado incrementando el flujo de presupuesto hacia el área de mantenimiento con el fin evitar que interesados malintencionados accedan por medio de formas no permitidas a sus datos y posteriormente efectúen daños industriales catastróficos como el cambio de parámetros en cierta maquina los cual generaría en una producción a escalas industriales perdidas económicas por producto defectuosos.

En contraposición a estos riesgos aparecen profesionales que conocen los riesgos y las mejores prácticas para mitigarlos utilizando metodología basada en el mantenimiento predictivo y la detección de anomalías; de esta forma se procura extender la confiabilidad y salud de la maquina con acciones anticipadas a cualquier tipo de falla; en este trabajo, proponemos utilizar las herramientas que nos ofrece el aprendizaje automático para formular un modelo con el cual podamos detectar datos que no se ajustan al promedio de los datos, los cuales posiblemente se generaron por algún tipo de alteración o fallo en la máquina. Tomando datos de mantenimiento industrial y analizando aspectos que afectan su normal funcionamiento se realizara una mejora corrigiendo y eliminando la causa de estos datos con acciones de control o reparación.

Revisión de literatura

Una búsqueda en Scopus base de datos muestra el creciente avance en investigación y desarrollo: se refieren a documentos de internet de las cosas 1350, 642 de ellos sobre mantenimiento, 475 documentos de machine learning, 269 documentos de seguridad de la red, 186 documentos de data privacidad, 64 documentos de Monitoreo.

Actualmente tenemos muchas variaciones de algoritmos inspirados en la naturaleza que no ayudan a optimizar encontrando mejores soluciones tales como optimización de enjambre de partículas, Optimización de llama de polilla, Algoritmo de optimización de ballenas, optimizador lobo gris, búsqueda del cuco entre otros.

Ciberseguridad

Los mecanismos de control de acceso generalmente están implementados pero su implementación es a menudo defectuosa, con una falsa sensación de seguridad que se viola fácilmente. A pesar de esto es prácticamente imposible evitar el acceso inapropiado

(ya que las personas con el acceso puede cometer delitos o, sin saberlo, proporcionar acceso a otros), medidas se puede tomar para identificar rápidamente los posibles casos de violaciones de privacidad.

Los algoritmos de detección de anomalías se pueden usar para monitorear el acceso a los datos y marcar variaciones de la norma, por ejemplo, identificar cuándo las personas (que pueden tener acceso permisos) acceder a datos en puntos de tiempo que son anómalos en comparación con sus historial de acceso previo y el historial de acceso de otros con roles similares en el mismo organización.

La pérdida de datos privados de las personas también puede conducir a ataques a organizaciones donde los individuos están empleados, con el posible robo de secretos comerciales e intelectuales propiedad. Esto eventualmente puede conducir a la pérdida de mercados y una reducción drástica en rentabilidad, si los competidores producen los mismos productos a un precio más bajo, ya que no tiene que pagar los costos de investigación sufridos por la organización original.

Detectar anomalías en los patrones de acceso, junto con las frecuencias y patrones de el correo electrónico y otras comunicaciones de los empleados pueden ayudar a descubrir el potencial pérdida de datos confidenciales.

Mantenimiento predictivo

Este tiene como objetivo reducir los costos de mantenimiento del presupuesto de una planta mediante la determinación del estado del equipo mediante la supervisión continua o periódica de las señales que describen su estado. Gracias a una detección temprana de fallas, se evitan fallas fatales y es posible programar tareas de mantenimiento cuando son más rentables. Las tecnologías de mantenimiento predictivo son clave para extender la vida útil del equipo, reducir los costos de mantenimiento y aumentar la explotación de activos. El análisis de vibraciones es una de las técnicas más efectivas en equipos rotativos de alta velocidad para evaluar su estado, detectar defectos y evitar fallas fatales. De la inspección visual realizada por expertos humanos, que es el método más utilizado, El mantenimiento predictivo ha evolucionado a métodos automatizados que utilizan técnicas avanzadas de procesamiento de señales y modelos de reconocimiento de patrones como redes neuronales, lógica difusa y modelado empírico y físico basado en datos. A medida que el equipo comienza a fallar, puede presentar síntomas que pueden revelarse si se utiliza una metodología adecuada para detectar precursores de falla.

Detección de anomalías

Una anomalía es una "variación de la norma". Muchos campos científicos y de ingeniería se basan en la suposición de que existen procesos o comportamientos en la naturaleza que siguen ciertas reglas o principios generales, lo que resulta en el estado de un sistema, manifestado en datos observables. A partir de los datos, debemos formular hipótesis sobre la naturaleza del proceso subyacente, que puede verificarse mediante la

observación de datos adicionales. Estas hipótesis describen el comportamiento normal de un sistema, suponiendo implícitamente que los datos utilizados para generar las hipótesis son típicos de El sistema en cierto sentido. Sin embargo, las variaciones de la norma pueden ocurrir en los procesos, por lo tanto, los sistemas también pueden existir en estados anormales, lo que lleva a valores de datos observables que son diferentes de los valores observados cuando no ocurren tales variaciones de proceso / estado. La tarea de detección de anomalías es descubrir tales variaciones (a partir de la norma) en los valores de datos observados y, por lo tanto, inferir las variaciones en el proceso subyacente. Un problema fundamental es que no existe una definición única y simple que nos permita evaluar cuán similares son dos puntos de datos y, por lo tanto, cuán diferente es un punto de datos de otros en el conjunto de datos.

Motor de detección de anomalías de aprendizaje automático

Un motor de detección de anomalías de aprendizaje automático (MLAD) utiliza datos de sensores como temperatura, flujo, vibración y otras variables (multidimensionales) para crear un patrón multidimensional, es decir, el estado de una máquina se "aprende" en condiciones de funcionamiento saludables. Al aprender el estado saludable, el motor determina si se debe activar una alerta o no bajo ninguna desviación.

El segundo paso es pronosticar la probabilidad de una avería futura o, en otras palabras, evaluar la RUL (vida útil residual) de una máquina. Esto proporciona información valiosa sobre reparaciones urgentes o sobre actividades de mantenimiento planificadas.

Finalmente, puede integrarlos en paneles o sistemas de alarma que serán monitoreados por equipos de mantenimiento.

El problema de detección de anomalías para series temporales se puede formular como encontrar puntos de datos atípicos en relación con alguna señal estándar o habitual. Nuestro enfoque debe ser desde la perspectiva de la máquina, como picos inesperados, cambios de nivel que apuntan al deterioro de la salud de una máquina.

Máquinas de vectores de regresión

Estos se define como un conjunto de algoritmos de aprendizaje supervisado; este tipo de métodos trata problemas de clasificación y regresión. Utilizando un entrenamiento de datos se etiqueta las clases para que el SMV sea capaz de crear un modelo para predecir la clase de una nueva muestra.

Este enfoque estima las funciones del modelo utilizando una función promedio ponderada localmente, donde los pesos están determinados por una función del núcleo elegida adecuadamente. Las funciones gaussianas a menudo se usan como núcleos, un enfoque

que se ha popularizado con las Máquinas de vectores de soporte (SVM) que también se han utilizado para problemas de clasificación.

Los SVM han surgido como los algoritmos de aprendizaje elegidos para muchos problemas de clasificación, y combinan varias ideas teóricamente importantes (como la regularización, las representaciones de espacios de características no lineales y los algoritmos de optimización de puntos interiores) capturados en implementaciones eficientes. Un SVM con funciones de núcleo gaussianas se asemeja a un modelo de red neuronal conocido como función de base radial (RBF), cuyos nodos aplican funciones gaussianas a las entradas netas y los parámetros de tales funciones se aprenden mediante variantes del algoritmo de descenso de gradiente.

Optimización por enjambre de partículas

Un algoritmo PSO utiliza un enjambre de soluciones posibles llamadas partículas. Estas partículas se mueven por el espacio de búsqueda de acuerdo a unas simples reglas matemáticas. El desplazamiento de cada partícula está relacionado con la mejor posición obtenida, así como de la mejor posición global hallada en toda la zona de búsqueda.

En su versión histórica, es un método iterativo colectivo, anárquico (en el sentido original del término), con énfasis en la cooperación; es parcialmente al azar y sin selección. El objetivo de los primeros capítulos será detallar estas características y formalizarlas para obtener un modelo explotable que sea particularmente efectivo para problemas fuertemente no lineales.

Inicialmente veremos por qué y cómo este modelo puede tratar de forma continua y optimizaciones heterogéneas (es decir, en las que algunas de las variables son continuas y otras discretas, posiblemente codificando aspectos combinatorios) de manera uniforme. Luego estudiaremos algunas alternativas. El objetivo aquí no es hacer una encuesta exhaustiva, sino trabajar en una selección de aquellos que ya han demostrado ser de interés, o parecen más prometedores. En otras palabras, su elección es necesariamente subjetiva. Examinaremos en particular las versiones conocidas como adaptativas, cuya forma "definitiva", llamada TRIBES, no requiere ningún parámetro distinto de los que definen el problema.

El diseño del modelo de detección de anomalías

El monitoreo de acondicionamiento de una máquina se realiza observando la medición de un sensor (p. Ej., Temperatura, vibración) e imponiéndole límites, es decir, en condiciones normales de funcionamiento, los valores de medición están limitados por un valor máximo y mínimo (similar a los gráficos de control) Cualquier desviación es los límites definidos envía una alarma. Esto a menudo se define generalmente como detección de anomalías.

Formulación matemática

La formulación matemática para modelar detección de anomalías se muestra a continuación.

$$X_t = \{X_1, X_2, X_3, \dots\} \quad (1)$$

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t \quad (2)$$

$$L^i X_t = X_{t-i} \quad (3)$$

$$X_t = \sum_{i=1}^p \varphi_i L^i X_t + \varepsilon_t \quad (4)$$

$$\varepsilon_t = X_t - \sum_{i=1}^p \varphi_i L^i X_t \quad (5)$$

$$\varepsilon_t = \left(1 - \sum_{i=1}^q \varphi_i L^i\right) X_t \quad (6)$$

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (7)$$

$$L^i \varepsilon_t = \varepsilon_{t-i} \quad (8)$$

$$X_t = \varepsilon_t + \sum_{i=1}^q \theta_i L^i \varepsilon_t \quad (9)$$

$$X_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t \quad (10)$$

$$X_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (11)$$

$$\left(1 - \sum_{i=1}^q \varphi_i L^i\right) X_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t \quad (12)$$

$$X_t = (1-L)^d X_t \quad (13)$$

$$\left(1 - \sum_{i=1}^q \varphi_i L^i\right) (1-L)^d X_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t \quad (14)$$

La serie temporal de datos dispersos está dada por la función (1). Las restricciones del (2) al (6) definen los términos del retraso de datos y como denota los errores pasados, restricciones del (7) al (12) definen el modelo auto-regresivo y los términos de la media móvil. Realizando el último reemplazo de la ecuación (12) por (13) obtenemos la generalización matemática (14).

Método de solución

-Todos los paquetes de software utilizados son paquetes de código abierto. El lenguaje utilizado es Python y paquetes asociados como pandas, matplotlib.

-Para la detección de anomalías, se utiliza un paquete de código abierto desarrollado por el equipo de Facebook llamado Prophet.

El conjunto de datos proviene del repositorio de la NASA. *"Conjunto de datos de rodamiento", Repositorio de datos de pronósticos Ames de la NASA, Centro de investigación Ames de la NASA, Moffett Field, CA*

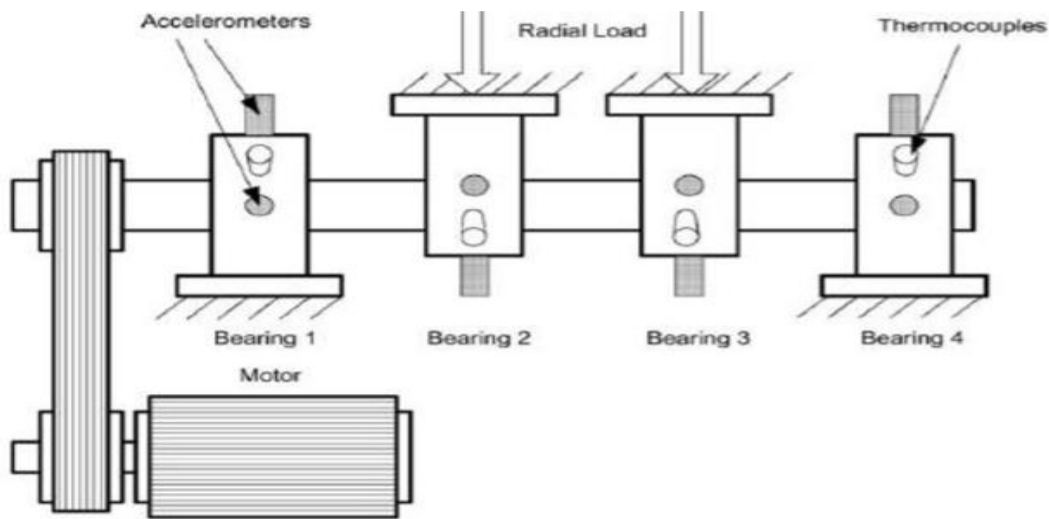


Fig 1. plataforma de prueba para la recopilación de datos

Importamos los paquetes de Python necesarios para el funcionamiento y procesamiento del modelo. Accedemos al punto de almacenamiento de datos y se cargan al modelo.

```
Archivo:
  creación de datos.zip : 2nd_test /
  inflado: 2nd_test / 2004.02.12.10.32.39
  inflado: 2nd_test / 2004.02.12.10.42.39
  inflado: 2nd_test / 2004.02.12.10.52.39
  inflado: 2nd_test / 2004.02.12.11.02.39
  inflado: 2nd_test / 2004.02.12.11.12.39
  inflado: 2nd_test / 2004.02.12.11.22.39
  inflado: 2nd_test / 2004.02.12.11.32.39
  inflado: 2nd_test / 2004.02.12.11.42.39
  inflado: 2nd_test / 2004.02.12.11.52.39
  inflado: 2nd_test / 2004.02.12.12.02.39
  inflado : 2nd_test / 2004.02.12.12.12.39
  inflado: 2nd_test / 2004.02.12.12.22.39
  inflado: 2nd_test / 2004.02.12.12.32.39
  inflado: 2nd_test / 2004.02.12.12.42.39
  inflado: 2nd_test / 2004.02.12.12.52.39
  inflado: 2nd_test / 2004.02.12.13.02.39
  inflado: 2nd_test / 2004.02.12.13.12.39
  inflado: 2nd_test / 2004.02.12.13.22.39
  inflado: 2nd_test / 2004.02.12.13.32.39
  inflado: 2nd_test / 2004.02.12.13.42.39
  inflar: 2nd_test / 2004.02.12.13.52.39
  inflar: 2nd_test / 2004.02.12.14.02.39
  inflar: 2nd_test / 2004.02.12.14.12.39
  inflar: 2nd_test / 2004.02.12.14.22.39
  inflado : 2nd_test / 2004.02.12.14.32.39
  inflado: 2nd_test / 2004.02.12.14.42.39
  inflado: 2nd_test / 2004.02.12.14.52.39
  inflado: 2nd_test / 2004.02.12.15.02.39
  inflado: 2nd_test / 2004.02.12.15.12.39
  inflado: 2nd_test / 2004.02.12.15.22.39
```

Fig 2.muestra de proceso de carga de datos

-Conjunto de datos

Los archivos de medición están disponibles como archivos separados (Total 983)

Inicio de los datos del Directorio

Datos / 2004.02.12.10.32.39 (12 de febrero de 2004)

Datos / 2004.02.12.10.42.39

Datos del directorio final

Datos de directorio final / 2004.02.19.06.12.39 (19 de febrero de 2004)

Datos / 2004.02.19.06.22.39

Cada archivo contiene 10 minutos de datos de acelerómetro.

-Lucha de datos

Cada uno de los archivos de 10 minutos se promedia y se almacena en un marco de datos de pandas

	Bearing 1	Bearing 2	Bearing 3	Bearing 4
2004-02-12 10:32:39	0.058333	0.071832	0.083242	0.043067
2004-02-12 10:42:39	0.058995	0.074006	0.084435	0.044541
2004-02-12 10:52:39	0.060236	0.074227	0.083926	0.044443
2004-02-12 11:02:39	0.061455	0.073844	0.084457	0.045081
2004-02-12 11:12:39	0.061361	0.075609	0.082837	0.045118
...
2004-02-19 05:42:39	0.453335	0.161016	0.137440	0.119047
2004-02-19 05:52:39	0.337583	0.132400	0.144992	0.092125
2004-02-19 06:02:39	0.351111	0.152266	0.151299	0.100817
2004-02-19 06:12:39	0.001857	0.003732	0.003656	0.001786
2004-02-19 06:22:39	0.001168	0.000767	0.000716	0.001699

983 rows × 4 columns

Fig 3.marco de datos de librería pandas

Visualizando datos de vibración

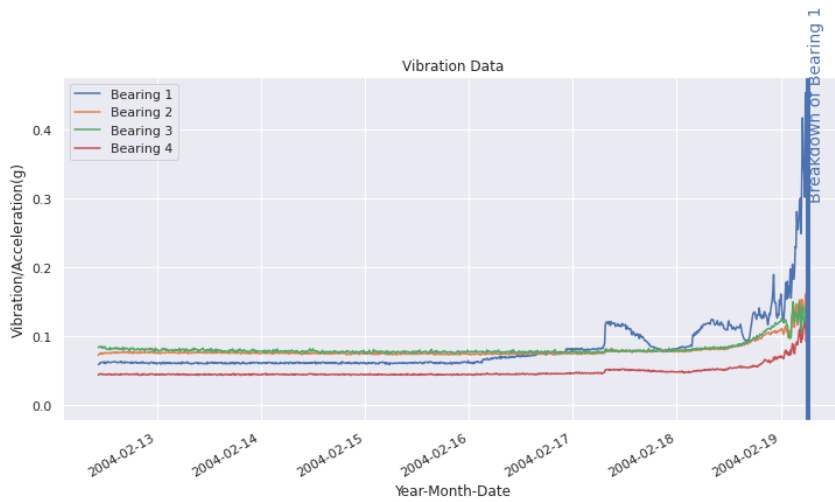


Fig 3.datos de vibración para los cuatros rodamientos.

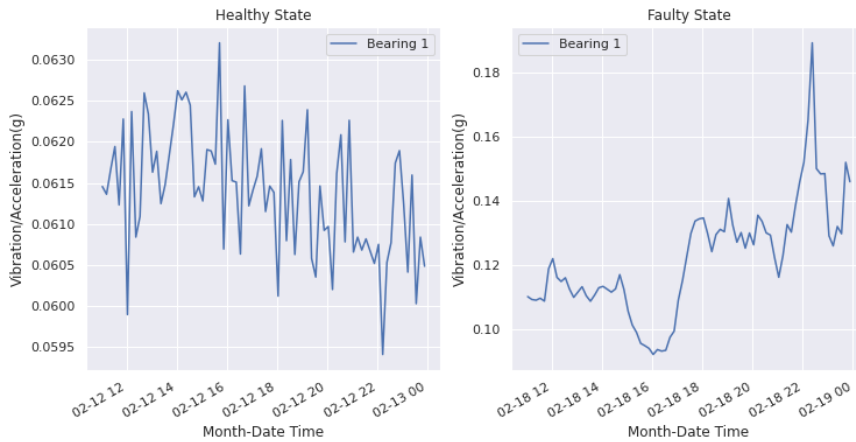


Fig 4. Datos de vibración del rodamiento 1 en sus dos estados. Grafica izquierda saludable y grafica derecha defectuosa.

Detección de anomalías usando Prophet Package

El problema de detección de anomalías para series temporales se puede formular como encontrar puntos de datos atípicos en relación con alguna señal estándar o habitual. Nuestro enfoque será desde la perspectiva de una máquina, como picos inesperados, cambios de nivel que apuntan al deterioro de la salud de una máquina.

Prophet es un software de código abierto lanzado por el equipo de Core Data Science de Facebook.

Conjunto de datos de entrenamiento

	ds	y
0	2004-02-12 10:32:39	0.058333
1	2004-02-12 10:42:39	0.058995
2	2004-02-12 10:52:39	0.060236
3	2004-02-12 11:02:39	0.061455
4	2004-02-12 11:12:39	0.061361

Fig 5. Datos de entrenamiento del modelo.

Evaluación de datos de capacitación

Posteriormente de entrenar los datos de la fase saludable se procede a evaluar el modelo creado

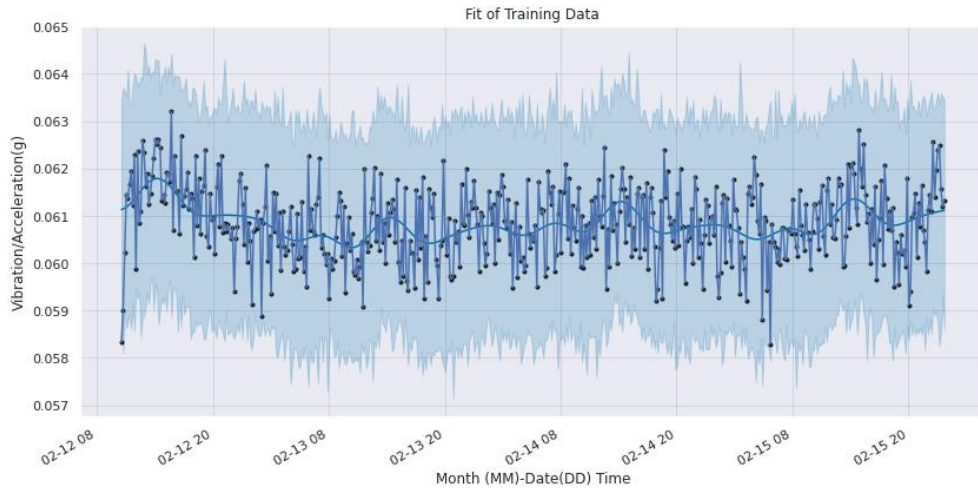


Fig 6. Ajuste de los datos de entrenamiento.

Los puntos negros son los puntos de datos del sensor de vibración. La línea azul representa la línea de tendencia con la porción azul claro que muestra la variación aceptable.

Evaluación de datos de falla.

En esta parte del proceso revisaremos si puede el modelo desarrollado anteriormente detectar una condición defectuosa de la máquina. Para evaluar esto, la porción de falla de los datos alimenta al modelo.

Tenga en cuenta que estos datos no han sido entrenados por el modelo de entrenamiento.

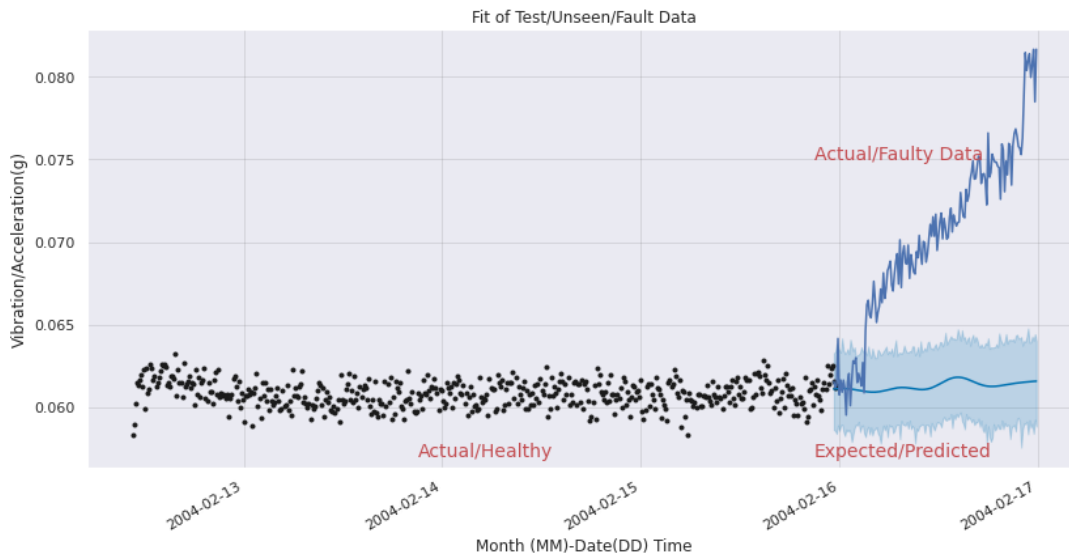


Fig 7. Ajuste de los datos de falla al modelo.

Los puntos negros son los verdaderos puntos de datos del sensor de vibración. La línea azul representa los valores esperados del 15/02/2004, 23:42:39 con la porción azul claro que muestra la varianza aceptable.

Claramente, los valores son más altos que los valores predichos y puede sonar una alarma.

Calculo de errores del modelo

Error porcentual absoluto medio (MAPE): 12.414082

Error porcentual medio (MPE): 12.339924

Error absoluto medio (MAE): 0.905349

Error cuadrático medio (RMSE): 0.010342

Optimización por enjambre de partículas

Utilizando el mismo conjunto de datos realizaremos un análisis que utiliza conductas y movimientos para generar un algoritmo que optimice las mejores posiciones o rango de datos.

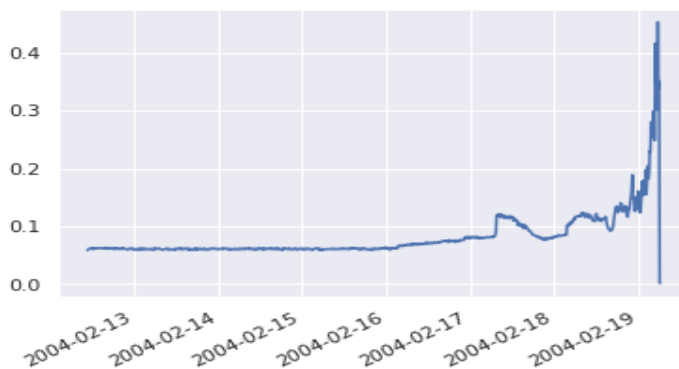


Fig 8. Grafica de datos en (PSO).

Para esta optimización por enjambre de partículas (psa) utilizamos 12 partículas, 10 iteraciones, 3 dimensiones, velocidad de detención de 1

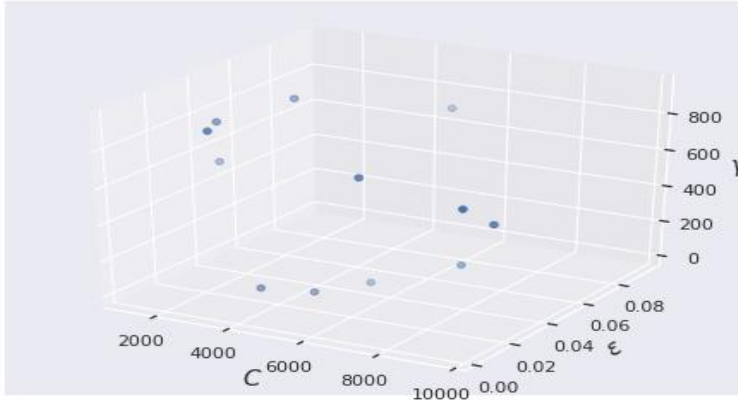


Fig 9. Comienzo del modelo con las partículas en posición aleatoria.

Mejor valor de iteración # 0 = 0.001956
 Mejor valor de iteración # 1 = 0.000114
 Mejor valor de iteración # 2 = 0.000074
 Mejor valor de iteración # 3 = 0.000066
 Mejor valor de iteración # 4 = 0.000066
 Mejor valor de iteración # 5 = 0.000065
 Mejor valor de iteración # 6 = 0.000065
 Mejor valor de iteración # 7 = 0.000064
 Mejor valor de iteración # 8 = 0.000064
 Mejor valor de iteración # 9 = 0.000064

Grupo Mejor configuración encontrada:
 [7.1697225e + 02 1.0000000e-08 1.0000000e-03]

Mejor regresor:

SVR (C = 716.9722501157323, cache_size = 200, coef0 = 0.0, grado = 3, epsilon = 1e-08,
 gamma = 0.001, kernel = 'rbf', max_iter = -1, encogimiento = Verdadero, tol = 0.001,
 detallado = Falso).

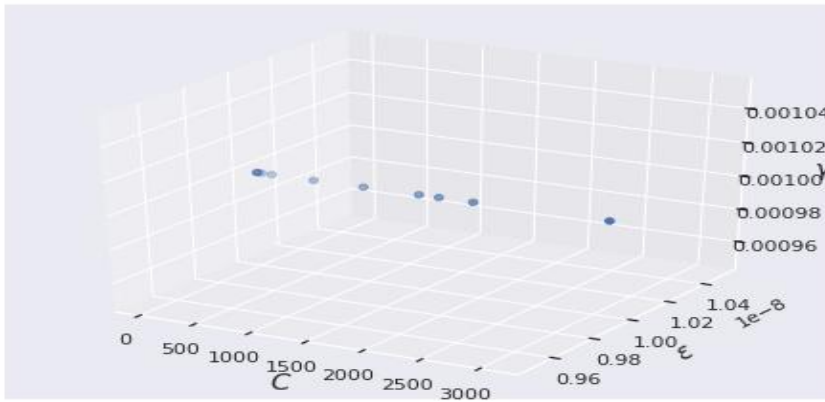
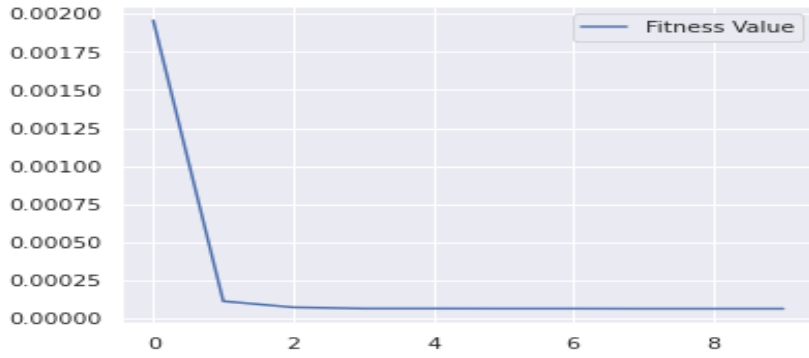


Fig 10. Proceso de convergencia de los datos a cero disminuyendo el error promedio.

Predicciones generadas

```

predicciones [0,24596606 0,24012802 0,24688443 0,25132741 0,25094931 0,24998393
0,25101696 0,25412693 0,25337904 0,25478196 0,25571911 0,25429979
0,25104923 0,26202526 0,26634648 0,265342 0,2658913 0,25913942
0,25229434 0,24858328 0,25221247 0,25384564 0,26245793 0,25209709
0,24851615 0,25315512 0,2479903 0,25809187 0,25542485 0,25089452
0,24996484 0,23961617 0,23680552 0,23489416 0,23441728 0,23573482
0,23381225 0,25517749 0,26194868 0,24951847 0,24680568 0,24925398
0,24174698 0,23629658 0,2398075 0,24335609 0,23732837 0,23374616
0,23783524 0,24264762 0,24357388 0,24169292 0,23974338 0,24189889
0,25131848 0,24156551 0,2271469 0,21798463 0,21318081 0,20589699
0,20432874 0,20253938 0,19851821 0,20178047 0,20075718 0,20129103
0,2098368 0,21395829 0,23455466 0,24743695 0,26315384 0,27856728
0,28676533 0,28822114 0,28882833 0,27849953 0,26655846 0,27784264
0,28129299 0,27972916 0,30172353 0,28416633 0,27278934 0,27914065
0,26883014 0,27885594 0,27103936 0,29064002 0,28660586 0,27895077
0,27748489 0,26259396 0,24958373 0,26395568 0,28442541 0,27943272
0,29711751 0,31258584 0,32557745 0,35358077 0,40462095 0,32144414
0,31793218 0,31814093 0,27702794 0,27033162 0,28316819 0,27838332
0,32556947 0,31277737 0,34472424 0,30645238 0,27790874 0,26541556
0,26747672 0,29995315 0,34473555 0,3798208 0,3286391 0,3841471
0,33352542 0,36789364 0,42195894 0,33229222 0,3490912 0,36923053
0,43654731 0,39170636 0,39023713 0,41668647 0,49184179 0,48734983
0,59858074 0,54242004 0,54489235 0,57509154 0,62795259 0,63986771
0,53092718 0,70216172 0,75049553 0,72144268 0,8879071 0,64305685
0,96558962 0,71999459 0,74871209 0,00678904]

```

Resultados

Para el desarrollo de este trabajo utilizamos dos herramientas tales como detección de anomalías con fb-prophet y la optimización de enjambre de partículas con los cuales se obtuvo predicciones muy cercanas a los valores reales lo cual genera una confiabilidad de los modelos muy alta. Sin embargo, se encuentra una diferencia cuando se haya el error cuadrático medio de los dos modelos utilizados.

-Detección de anomalías con fb-prophet: 0.010342

-Optimización de enjambre de partículas: 0.006448

Comprobación de ajuste de predicciones generadas en PSO con los datos reales

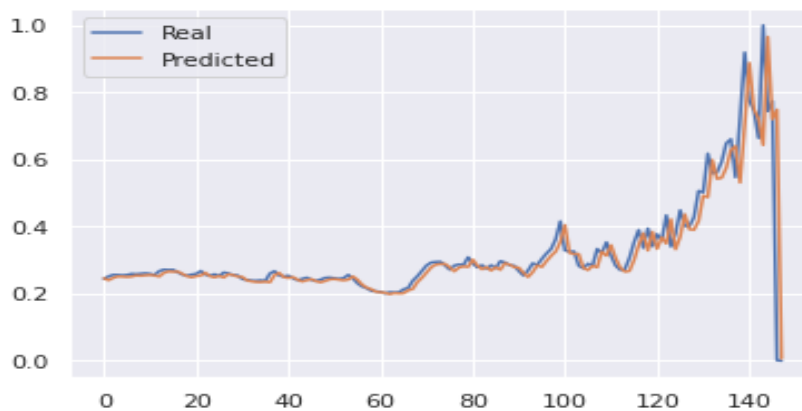


Fig 11. Predicciones con el mejor valor de población encontrado.

Conclusiones

-Por medio del avance en tecnologías y el internet de las cosas se ha activado más el robo de información y ataques a maquinaria hacia empresas por interesados malintencionados

-Las gráficas generadas en este trabajo muestran el avance del mantenimiento predictivo ofreciendo datos con los cuales la programación de mantenimiento evitara millones en costos.

-Usando herramientas como las utilizadas en este trabajo se está dando cumplimiento con la normativa de la ISO 27001 al proteger y controlar la información y datos como las fallas de maquinaria industrial.

-El funcionamiento de los ataques en ciber seguridad se basan robar, copiar y destruir datos, aprovechando el descuido y poco mantenimiento de máquinas y sistemas vitales para una empresa.

- El modelo entrenado de fb-prophet predice las tendencias para el futuro y si se observa alguna desviación, se puede generar una alarma. El mismo principio es válido para analizar múltiples señales (multidimensionales) a la vez y crear una sola métrica como el "puntaje de salud" de una máquina.

-Usando estrategias de mantenimiento predictivo, uno puede encontrar las primeras etapas de daño y reducir los costos operativos. En lugar de apagar el equipo para las rutinas de mantenimiento preventivo programado

-El uso del Internet de las cosas industrial (IIoT) permite el acceso a datos en tiempo real y permite construir los análisis estadísticos y de tendencias que señalan por qué y dónde ocurren los problemas.

Bibliografía

- Abraham B. and Chuang A. Outlier detection and time series modeling. *Technometrics* 31(2), 241–248 (1989).
- Agarwal D. Detecting anomalies in cross-classified streams: A bayesian approach. *Knowledge Information Systems* 11(1), 29–44 (2006).
- ANGELINE P.J., “Using Selection to Improve Particle Swarm Optimization”, IEEE International Conference on Evolutionary Computation, Anchorage, Alaska, May 4-9, p. 84-89, 1998.
- B. Schölkopf, A. J. Smola *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning)*, MIT Press, 2002.
- BERTELLE C., DUTOT A., GUINAND F., Olivier D., “Color ant populations algorithm for dynamic distribution in simulations”, poster presented at European Simulation Symposium and Exhibition, Delft, the Netherlands, 2003.
- C. Cortes, V. Vapnik, “Support-vector networks.” *Mach. Learn.* 20(3), 273–297 (1995)
- C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121-167, 1998.
- C.-C. Chang, C.-W. Hsu, and C.-J. Lin. The analysis of decomposition methods for support vector machines. *IEEE Transactions on Neural Networks*, 11(4):1003-1008, 2000.

- C.C.Aggarwal,C.K.Reddy,DataClustering:AlgorithmsandApplications(CRCPress,Bo ca Raton, 2013)
- CARLISLE A., DOZIER G., "Adapting Particle Swarm Optimization to Dynamics Environments", International Conference on Artificial Intelligence, Las Vegas, Nevada, USA, p. 429-434, 2000.
- CLERC M., "The Swarm and the Queen: Towards a Deterministic and Adaptive Particle Swarm Optimization", Congress on Evolutionary Computation, Washington D.C., p. 1951-1955, 1999.
- G. Hale, "InTech survey: Predictive maintenance top technology challenge in 2007," InTech2007, Jan. 2007, (ISA: International Society of Automation).
- H.M. Hashemian, "State-of-the-art Predictive Maintenance techniques," IEEE Transactions on Instrumentation and Measurement, vol. 60, No. 1, pp. 226-236, 2011.
- J. Han, M. Kamber, J. Pei, "Data Mining,(2006): Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems)" .
- J. Lee, H. Qiu, G. Yu, J. Lin (2007). IMS, Universidad de Cincinnati. "Conjunto de datos de rodamiento", Repositorio de datos de pronósticos Ames de la NASA
- Karthick Thiyagarajan, Sarath Kodagoda, Linh Van Nguyen, "Predictive Analytics for Detecting Sensor Failure Using Autoregressive Integrated Moving Average Model," 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA).
- Kishan G, (2017). Anomaly Detection Principles and Algorithms
- Kumpati S. Narendra, Kannan Parthasarathy (1990). "Identification and Control of Dynamical Systems Using Neural Networks". IEEE Transactions On Neural Networks. Vol. I . N°. I. March 1990
- O. Bousquet and A. Elisseeff. Stability and generalization. Journal of Machine Learning Research, 2001. Submitted.
- P. Mondal, L. Shit, and S. Goswami, "Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices," International Journal of Computer Science, Engineering and Applications, vol. 4, p. 13, 2014.
- P.K.Kankar,C.S.Sharma,S.P.Harsha,"Fault Diagnosisofballbearings using machine learning methods", Expert Systems with applications, vol. 38, pp. 1876-1886, 2011.
- R. Fujimaki, T. Yairi, K. Machida, "An anomaly detection method for spacecraft using relevance vector," in Learning, The Ninth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) (Springer, New York, 2005), pp. 785–790
- R. K. Mobley, An Introduction to Predictive Maintenance, 2nd ed. St. Louis, MO: Butterworth-Heinemann, 2002.
- Ted Dunning, (2014). Practical Machine Learning: A New Look at Anomaly Detection