Artículo de Investigación

Uso de algoritmo K-media para detectar estado de funcionamiento de electroválvula

K-Means algorithm to detect solenoid valve function status.

Yánez, Diego; Molina, Iván; Gallardo, Andrés

Diego Yánez diego_alex40@hotmail.com Vinnytsia National Technical University, Ucrania Iván Molina ivanrm10q@hotmail.com Escuela Politécnica Nacional, Italia

Andrés Gallardo andressebastian.gallardo@mail.polimi.it Student at Politecnico di Milano, Italia

Ecuadorian Science Journal GDEON, Ecuador ISSN-e: 2602-8077 Periodicidad: Semestral vol. 5, núm. Esp.3, 2021 esj@gdeon.org

Recepción: 31 Agosto 2021 Aprobación: 04 Octubre 2021

URL: http://portal.amelica.org/ameli/ jatsRepo/606/6062738020/index.html

DOI: https://doi.org/10.46480/esj.5.3.157

Los autores mantienen los derechos sobre los artículos y por tanto son libres de compartir, copiar, distribuir, ejecutar y comunicar públicamente la obra sus sitios web personales o en depósitos institucionales, después de su publicación en esta revista, siempre y cuando proporcionen información bibliográfica que acredite su publicación en esta revista. Licencia de Creative Commons Las obras están bajo una https://creativecommons.org/licenses/by-ncnd/4.0/deed.es



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivar 4.0 Internacional.

Como citar: Yánez , D., Molina , I., & Gallardo , A. (2021). Uso de algoritmo K-media para detectar estado de funcionamiento de electroválvula. Ecuadorian Science Journal, 5(3), 220-238. DOI: https://doi.org/10.46480/esj.5.3.157



Resumen: En la actualidad las aplicaciones de la inteligencia artificial dentro de los procesos de ingeniería en general son extremadamente amplios, dado las características que tiene tanto la Inteligencia artificial como la Ingeniaría en general. En este trabajo investigativo exponemos el uso de la inteligencia artificial, en específico el agrupamiento o "clustering" para poder detectar cuando una válvula se encuentra con fallas o no, de esta forma poder conocer, de manera precisa, el estado de funcionamiento de la electroválvula por Inteligencia Artificial a partir de datos recopilados, dotándole del análisis realizado en Matlab como Data-Set para su proceso de entrenamiento. El elemento de estudio de este artículo es una electroválvula de tres posiciones con compensación de resorte (DIVW estilo C). La válvula se utiliza en el presente experimento en dos regiones de funcionamiento diferentes: a 5 MPa y 20 MPa. Además, se agregó al análisis un escenario en el que la válvula presenta un daño, mismo que consiste en una hendidura de 0,1 [mm] de profundidad.

Palabras clave: Automatización, Inteligencia Artificial, Tecnología Electrónica.

Abstract: Now days, the applications of artificial intelligence within engineering processes in general are extremely broad, given the characteristics of both artificial intelligence and engineering in general. In this investigative work we expose the use of artificial intelligence, specifically clustering to be able to detect when a valve is faulty or not, in this way to be able to know, precisely, the operating state of the solenoid valve by Artificial Intelligence from collected data, providing with the analysis carried out in MatLab as a Data-Set for training process. The item of study for this article is a spring compensated three position solenoid valve (DIVW style C). The valve is used in the present experiment in two different operating regions: at 5 MPa and 20 MPa. In addition, a scenario in which the valve presents damage was added to the analysis, which consists of a groove 0.1 [mm] deep.

Keywords: Automation, Artificial Intelligence, Electronic Technology.



Modelo de publicación sin fines de lucro para conservar la naturaleza académica y abierta de la comunicación científica Diego Yánez, et al. Uso de algoritmo K-media para detectar estado de funcionamiento de electroválv...

INTRODUCCIÓN

Este proyecto se basa en el análisis del comportamiento de una válvula de presión utilizada para maquinarias de fundición a presión. El objetivo de este estudio es identificar adecuadamente el régimen de trabajo de la válvula a través de medidas de las vibraciones y la corriente de los componentes de la válvula, los cuales eran procesados por una Inteligencia artificial. La medición de las vibraciones se realiza mediante un acelerómetro acoplado a la válvula, mientras que la medición de la corriente se realiza mediante un sensor de corriente que mide la corriente del solenoide dichos datos se recopilaran sobre dos diferentes tipos de funcionamiento 5 MPa y 20 MPa. Además, se agregó al análisis un escenario en el que la válvula presente un daño, mismo que consiste en una hendidura de 0,1 [mm] de profundidad en el carrete de la válvula, en la parte que conecta los canales A y B (figura 1). En consecuencia, tenemos cuatro condiciones de trabajo de la válvula en la que se adquirieron los datos.



FIGURA 1. Daño del carrete de la válvula. Experimento Prof. Marco Tarabini & Prof. Fabio Conti

El ciclo de funcionamiento de la válvula presenta dos secuencias o fases principales, las cuales se dividen en A y B tal y como podemos verificar en la figura 2.



FIGURA 2.

(Izquierda) Posición cerrada de la válvula. (Derecha) Posiciones abiertas de la válvula; fase A y fase C. Bulletin HY14-2536-M1/US, 600, 4/07(Parker Corporation, 2007)

La válvula opera de manera continua de la siguiente manera: Fase A (a) -cerrada (b) -Fase C (c) -cerrada (d), durante doce minutos y treinta y seis segundos. La Figura 3 muestra este punto, y la clara manera de poder diferenciar los diferentes momentos de la fase y de como recopilaron los datos.



Aceleración vs. tiempo en la fase A y C durante el ciclo de operación. Autoría Propia

Procesamiento de datos

Después de la correcta medición se creará una inteligencia artificial usando el método de K-media a fin de poder identificar de manera automática los estados de daño lo cual crea cuatro posibles variantes en los que se puede clasificar la información recopilada. (García & Gómez, 2006). Se escogió el tipo K-media, dado que generalmente en el agrupamiento se comienzan a calcular las distancias a partir de un número predefinido de grupos y se van colocando de forma iterativa los genes en los diferentes grupos hasta minimizar la dispersión interna de cada uno, lo cual se conoce como algoritmo no jerárquico (Milone et al., 2009).

A fin de poder obtener un proceso investigativo adecuado se ha escogido metodología cuantitativa dada la necesidad de medir un fenómeno en concreto y el proceso de recolección de datos (Hernández Sampieri & Mendoza Torres, 2018).De esta manera se forma un Data-Set que cumpla con los requerimientos necesarios para poder dotar a la Inteligencia Artificial en su proceso de aprendizaje.

Con el fin de obtener información útil a partir de la cual sea posible iniciar el análisis, es importante transformar los datos brutos de los instrumentos de medición en valores con las unidades de ingeniería adecuadas de los fenómenos que estamos midiendo.

Para este caso tenemos datos provenientes de un acelerómetro y de un sensor de corriente. Ambos dispositivos dan valores de salida de voltaje en un intervalo de tiempo específico. Para transformar los valores del acelerómetro en unidades de ingeniería, usamos el parámetro llamado sensibilidad. El acelerómetro utilizado es el Delta Tron Type 4397 producido por la empresa Brüel & Kjaer ytiene una sensibilidad de 0,991 mV/ms2. Por lo tanto, se divide cada voltaje del acelerómetro por la sensibilidad dado en la ecuación [1] de la siguiente manera.

$$Aceleracion = \frac{Voltaje_{salida}(datos brutos)}{Sensibilidad}$$
[1]

Usando la relación anterior, fue posible construir las Figuras 4 y 5 de aceleración versus tiempo en cada uno de los cuatro estados posibles para la válvula de cada caso, dejando una clara diferencia entre los posibles estados.



FIGURA 4. Aceleración vs. tiempo del caso sano de la fase C cuando la válvula está funcionando a 5 MPa. _{Autoría Propia}



FIGURA 5

Aceleración vs. tiempo del caso dañado de la fase C cuando la válvula está funcionando a 5 MPa. _{Autoría Propia}

Al realizar la graficas se puede identificar que, en el caso de 5 MPa, cuando la válvula está sana (figura 4) y dañada (figura 5) obtenemos datos diferentes lo cual se representa gráficamente sin problema alguno, notando un abultamiento gráfico en torno a la línea de cero (0), en la aceleración cuando se encuentra funcionando correctamente. Mismo caso se encuentra en el caso de 20 MPa, lo cual se puede apreciar al comparar las figuras 6 y 7.



FIGURA 6 Aceleración vs. tiempo del caso sano de la fase A cuando la válvula está funcionando a 20 MPa. _{Autoría Propia}



FIGURA 7.

Aceleración vs. tiempo del caso dañado de la fase C cuando la válvula está funcionando a 20 MPa. _{Autoría Propia}

Para resumir, hay cuatro escenarios de adquisición de datos, mismo que son tanto "saludables" como "dañados" a una presión de trabajo de la válvula a 5 MPa y a 20 MPa.

Análisis de datos brutos

Una vez obtenidos estos gráficos fue posible analizar los fenómenos, con lo que se identificó un comportamiento similar cada 6 segundos cuando se acercaban los gráficos. Este razonamiento sugirió considerar solo un intervalo de tiempo dentro de los 6 segundos de cada ciclo.

Sin embargo, era importante tomar en cuenta un número mínimo de ciclos para asegurarse que se desarrolló un comportamiento similar en todos ellos. En cuanto a la aceleración, no hubo ningún comportamiento particular en todos los ciclos del escenario de 5 MPa, ni en el caso sano ni en el caso defectuoso. Contrariamente, en el escenario de 20 MPa, se identificó algunos ciclos, especialmente en el caso defectuoso, donde la curva adquiría una forma particular, diferente a todos los demás casos que teníamos. Este extraño camino se puede ver en la figura 8, resaltados en círculos de color rojo.



FIGURA 8.

Aceleración vs. tiempo del caso sano y defectuoso a una presión de trabajo de 20 MPa (Fase A). Los ciclos con el error debido a la saturación del acelerómetro se identifican con círculos rojos. Autoría Propia

Este fenómeno se conoce como "zero-shift" y provoca la saturación del acelerómetro debido al alto g-choque alcanzado durante la prueba a 20 MPa (Song & Huang, 2005). Se eliminaron los ciclos que presentaban este error para evitar la influencia del mismo en el análisis. Además, se observó que el error de compensación debido a la saturación del acelerómetro fue menor al final de los ciclos.

Redefinición de intervalo: Como se argumentó anteriormente, se identificaron ciclos de 6 segundos cada uno para cada fase, para corriente y aceleración. Entonces, fue factible analizar solo unos pocos ciclos para comprender los fenómenos de una manera más eficiente. Dentro de la duración de cada ciclo (6 segundos) entendimos que se podría tomar un intervalo más corto para fines de agrupamiento. Las siguientes figuras 9 y 10 nos ayudarán a comprender mejor cómo elegimos el intervalo para cada ciclo. Las figuras 9 y 10 muestran que al inicio y al final de cada ciclo tendremos valores con amplitudes realmente altas que no muestran un comportamiento estable, por lo que consideramos elegir un intervalo entre 1 y 5 segundos.



FIGURA 9. Aceleración vs. tiempo de casos sanos y defectuosos para una presión de 5 MPa (Fase A). _{Autoría Propia}



FIGURA 10. Aceleración vs. tiempo de casos sanos y defectuosos para una presión de 5 MPa (Fase A). _{Autoría Propia}

En las figuras donde se analizó las diferentes características, era bastante factible seleccionar un intervalo entre los segundos 2 y 5 de cada ciclo. El intervalo de tiempo elegido para este análisis fue entre los segundos 4 y 5 (1 segundo). Este intervalo se seleccionó para optimizar el tiempo de cálculo. Además, observamos que el error de compensación debido a la saturación del acelerómetro fue menor al final de los ciclos, por lo que, si hubiera un error debido a la saturación del acelerómetro, su influencia se reduciría.

Análisis de funciones en el dominio del tiempo: Para el análisis del dominio del tiempo, se extrajeron las características de la aceleración y los datos actuales. Las fórmulas X, X y X fueron utilizadas tanto para la aceleración como para la corriente. A fin de buscar la mejor opción para la carga de datos se verificó y graficó los siguientes parámetros: Valor medio en relación a la Aceleración (ecuación 2), Desviación estándar (3), Raíz Cuadrada Media (RMS) (ecuación 4), Oblicuidad (Skewness) (ecuación 5), Curtosis (ecuación 6) y Pico a Pico (Peak to Peak) (ecuación 7)

Valor medio: Responde a la ecuación 2.

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{2}$$

Donde:

 μ = media n = número de elementos x = aceleración / corriente

Con esta fórmula se realizó una representación gráfica, tal como se puede ver en las figuras 11 y 12, a fin de entender el comportamiento y escoger los datos que van a alimentar a la IA.



Valor medio para cada segundo a 5 MPa - fase A y Valor medio de cada segundo a 20 MPa - fase A (Aceleración) _{Autoría Propia}

En la comparación de estas graficas no se encuentra una distinción clara entre la lectura de los casos de falla y saludables, lo cual lo hace poco elegibles para formar un parte de la data set principal que se debe cargar a la Inteligencia Artificial.

Desviación estándar: Responde a la ecuación 3 y con la misma se realizaron las figuras 13 y 14, correspondientes a 5 MPa y 20Mpa cada una, de esta manera exponiendo así su relación atreves del tiempo.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2}$$
^[3]

Donde:

 $\mu = media$

 σ = desviación estándar

n = número de elementos

x = aceleración / corriente



FIGURA 12 Desviación estándar de cada segundo a 5 MPa - fase A y Desviación estándar de cada segundo a 20 MPa - fase A (Aceleración) _{Autoría Propia}

Raíz Cuadrada Media (RMS)

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i)^2}$$
^[4]

Donde: RMS = raíz cuadrada media n = número de elementos x = aceleración / corriente



FIGURA 13. Raíz cuadrada media de cada segundo a 5MPa y 20 MPa - fase A (Aceleración) _{Autoría Propia}

Donde:

- μ = media S = asimetría M3 = tercer momento central σ = desviación estándar n = número de elementos
- x = aceleración / corriente



FIGURA 14. Oblicuidad (Skewness) de cada segundo a 5 MPa y 20 MPa - fase A (Aceleración) _{Autoría Propia}

Curtosis

$$K = \frac{M_4}{\sigma^4} - 3; \ M_4 = \frac{1}{n} \sum_{i:1}^n [(x_i - \mu)^4]$$
[6]

Donde:

 μ = media K = curtosis M4 = cuarto momento central σ = desviación estándar n = número de elementos x = aceleración / corriente



FIGURA 15 Curtosis de cada segundo a 5MPa y 20 MPa - fase A (Aceleración) _{Autoría Propia}

Pico a Pico (Peak to Peak)

$$Pk - Pk = X_{max} - X_{min}$$
^[7]

Donde: Pk-Pk = pico a pico X_max = valor máximo de la señal X_min = valor mínimo de la señal



FIGURA 16 Pico a Pico de cada segundo a 5MPa y 20 MPa - fase A (Aceleración) _{Autoría Propia}

Aplicación de medidas a corriente: A fin de llegar a una mejor conclusión se realiza una aplicación gráfica con una aplicación a la corriente en lugar de la aceleración dando como resultado los siguientes gráficos 18, 19, 20, 21 y 22:





FIGURA 19. Raíz cuadrada media de cada segundo a 5MPa y 20 MPa - fase A (Corriente) _{Autoría Propia}



FIGURA 20.

Oblicuidad (Skewness) de cada segundo a 5 MPa y 20 MPa - fase A (Corriente) Autoría Propia



FIGURA 21. Curtosis de cada segundo a 5MPa y 20 MPa - fase A (Corriente) _{Autoría Propia}



Pico a Pico de cada segundo a 5MPa y 20 MPa - fase A (Corriente) Autoría Propia

Dada la clara separación entre las señales saludables y con falla dentro de las medidas de pico a pico y de curtosis, se las escogió como los datos para armar el data set para su ingreso en la fase A y la fase C. Los datos tienen características útiles en el dominio del tiempo, por tanto, su agrupación se realizó en ambas fases, presentando en ambos, buenos resultados.

Agrupación para K-media

La agrupación en clústeres es una herramienta poderosa para procesar datos. Es un método de aprendizaje automático no supervisado, dado que no existe un maestro o guía para el proceso de aprendizaje(Russell & Norvig, 2007), lo cual es realmente útil cuando es necesario identificar estructuras en un conjunto de datos. Para el propósito de este proyecto, se utilizó la técnica de agrupamiento para identificar los cuatro escenarios (en los que la válvula de presión estaba funcionando) mencionados anteriormente. El enfoque de se puede visualizar con claridad en la figura 23.



L

La agrupación del conjunto de datos que o "clustering" se entiende como el proceso de agrupar datos en clases o "clusters" de tal forma que los objetos de un cluster tengan una similitud alta entre ellos, y baja (sean muy diferentes) con objetos de otros clusters (García & Gómez, 2006). En pocas palabras, este algoritmo selecciona centroides aleatorios y mide la distancia a cada punto, luego, realiza cálculos iterativos para mejorar la posición de los centroides.

Por lo tanto, se seleccionaron dos características con el fin de diferenciar los cuatro escenarios o clusters que se suponía iban a ser identificados. Una característica diferenciaría entre los casos de 5 MPa y 20 MPa; y la otra característica los casos sanos y defectuosos. Se notó, durante la extracción de las características, que la curtosis de la corriente era considerablemente diferente en un cierto intervalo de tiempo de cada ciclo para el caso sano y defectuoso. Este fenómeno se puede apreciar en la figura 24.



FIGURA 24 Curtosis de la corriente en el dominio del tiempo de los datos correspondientes a 5MPa (Fase A). _{Autoría Propia}

Este gráfico fue útil para comprender que los intervalos encerrados en verde de cada ciclo no fueron útiles, ya que las fechas de ambos casos coinciden en esos intervalos. Sin embargo, los datos entre los segundos 2 y 5 de cada ciclo fueron claramente diferentes para caso sano y defectuoso. Como se dijo anteriormente, para las características que se utilizarán en la agrupación, seleccionamos el intervalo entre los segundos 4 y 5. Los

valores de la característica de curtosis para el intervalo de 4-5 segundos para cada ciclo de la fase A se pueden ver gráficamente en la Figura XX



FIGURA 25 Valores de la curtosis de la corriente en el intervalo entre 4-5 segundos de cada ciclo de la fase A cuando la presión de trabajo es de 5 MPa. Autoría Propia

La agrupación en clústeres es una técnica ampliamente utilizada en informática con aplicaciones para el aprendizaje no supervisado, la clasificación, la minería de datos y otros campos.(Har-peled, 2004),

La función de agrupación de k-media de MATLAB requiere una matriz con las características y el número de agrupaciones deseadas como entradas. La tabla 1 explica cómo se estructuraron los datos de las características seleccionadas para la agrupación en la matriz necesaria para la función de MATLAB de k-medias. Si bien el proceso de K-media dentro de MATLAB no puede procesar imágenes (Al-Azzeh et al., 2019), es muy bueno en la identificación de clusters planos como es este caso. Los datos fueron introducidos de según se muestra en la Tabla1

$\begin{array}{c c c c c c c c c c c c c c c c c c c $			
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		Feature 1	Feature 2
$\begin{array}{c cccc} & \text{of the} & \text{of the} & \text{acceleration} & \text{in time} & \text{domain}) \\ \hline & \text{fume} & \text{domain}) & \text{domain}) \\ \hline & \text{S MPa.} & \text{K-h-51} & \text{FK2PK-h-51} & \text{FK2PK-h-52} & \text{FK2PK-h-53} & \dots & \text{FK2PK-f-51} & \text{FK2PK-f-51} & \text{FK2PK-f-51} & \text{FK2PK-f-53} & \dots & $		(Kurtosis	(Peak to peak
$\begin{array}{c} \begin{array}{c} \mbox{current in time domain} & \mbox{acceleration in time domain} \\ \mbox{f} & f$		of the	of the
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$		current in	acceleration
$\begin{array}{c c c c c c c c c c c c c c c c c c c $		time	in time
$ \begin{array}{c} {}^{5} {\rm MPa.} \\ {\rm Healthy} \\ {\rm Data} \\ \end{array} \begin{array}{c} {\rm K-h-51} \\ {\rm K-h-52} \\ {\rm K-h-53} \\ {\rm K-h-51} \\ {\rm K-f-51} \\ {\rm K-f-51} \\ {\rm K-f-52} \\ {\rm K-f-51} \\ {\rm K-f-52} \\ {\rm K-f-51} \\ {\rm K-f-52} \\ {\rm K-f-53} \\ {\rm K-f-53} \\ {\rm K-f-53} \\ {\rm K-f-53} \\ {\rm K-h-201} \\ {\rm K-h-201} \\ {\rm K-h-202} \\ {\rm K-h-203} \\ {\rm K-h$		domain)	domain)
5 MPa. $K-h-52$ $PK2PK-h-52$ Healthy $K-h-52$ $PK2PK-h-53$ Data $K-h-53$ $PK2PK-h-53$ 5 MPa. $K-f-51$ $PK2PK-f-51$ Faulty $K-f-52$ $PK2PK-f-53$ Data $K-f-53$ $PK2PK-f-53$ Data $K-f-53$ $PK2PK-f-53$ Data $K-f-53$ $PK2PK-f-53$ PMa. $K-f-53$ $PK2PK-f-53$ PMa. $K-f-201$ $PK2PK-h-201$ MPa. $K-h-203$ $PK2PK-h-203$ MPa. $K-h-20n$ $PK2PK-h-203$ Data $K-h-201$ $PK2PK-h-203$ $MPa.$ $K-f-201$ $PK2PK-h-203$ $MPa.$ $K-f-201$ $PK2PK-f-201$ 20 $K-f-201$ $PK2PK-f-201$ $MPa.$ $K-f-203$ $PK2PK-f-201$ $PK2PK-f-203$ $PK2PK-f-203$ $PK2PK-f-203$ $PK2PK-f-203$	5 MPa.	V_b_51	PK2PK-h-51
Healthy Data K-f-53 K-h-53 K-h-5n PK2PK-h-53 PK2PK-h-5n 5 MPa. Faulty Data K-f-51 K-f-52 K-f-53 K-f-53 K-f-5n PK2PK-f-51 PK2PK-f-52 PK2PK-f-53 20 MPa. Healthy Data K-h-201 K-h-202 K-h-203 K-h-203 PK2PK-h-202 PK2PK-h-201 20 MPa. Healthy Data K-h-201 K-h-203 PK2PK-h-202 PK2PK-h-203 20 MPa. Healthy Data K-h-201 K-h-203 PK2PK-h-202 PK2PK-h-203 20 MPa. Healthy Data K-f-201 K-h-203 PK2PK-h-203 20 MPa. K-f-202 MPa. Faulty K-f-201 K-f-203 PK2PK-f-203		K-II-51 K-h-52	PK2PK-h-52
Data K-fi-53 PK2PK-h-5n 5 MPa. K-f-51 PK2PK-f-51 Faulty K-f-52 PK2PK-f-52 Data K-f-53 PK2PK-f-53 Data K-f-50 PK2PK-f-53 Vata K-f-50 PK2PK-f-53 Vata K-f-201 PK2PK-h-201 VAPa. K-h-203 PK2PK-h-202 MPa. K-h-203 PK2PK-h-203 Data K-h-201 PK2PK-h-201 VAPA. K-h-203 PK2PK-h-202 MPa. K-h-201 PK2PK-h-203 MPa. K-f-201 PK2PK-f-201 20 K-f-201 PK2PK-f-201 MPa. K-f-203 PK2PK-f-202 MPa. K-f-203 PK2PK-f-202	Healthy	K-11-52	PK2PK-h-53
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	Data	N-11-55	
$ \begin{array}{c} 5 \text{ MPa.} \\ Faulty \\ Data \end{array} \\ \begin{array}{c} K-f-51 \\ K-f-52 \\ K-f-53 \\ K-f-5n \end{array} \\ \begin{array}{c} PK2PK-f-51 \\ PK2PK-f-52 \\ PK2PK-f-53 \\ PK2PK-f-5n \end{array} \\ \begin{array}{c} FK2PK-f-5n \\ PK2PK-f-5n \\ PK2PK-h-201 \\ PK2PK-h-201 \\ PK2PK-h-202 \\ F-h-203 \\ PK2PK-h-203 \\ PK2PK-h-203 \\ PK2PK-h-201 \\ PK2PK-h-201 \\ PK2PK-h-201 \\ PK2PK-f-201 \\ 0 \\ 0 \\ PK2PK-f-201 \\ PK2PK-f-201 \\ PK2PK-f-202 \\ PK2PK-f-202 \\ PK2PK-f-202 \\ PK2PK-f-202 \\ PK2PK-f-203 \\ PK2PK-f-$. K-II-DII	PK2PK-h-5n
$ \begin{array}{c} 5 \text{ MPa.} \\ Faulty \\ Data \\ 20 \\ MPa. \\ Healthy \\ Data \\ \end{array} \begin{array}{c} K-h-201 \\ K-h-202 \\ K-h-203 \\ K-h-203 \\ K-h-203 \\ K-h-201 \\ K-h-203 \\ K-h-203 \\ K-h-203 \\ K-h-201 \\ K-h-203 \\ K-h-203 \\ K-h-201 \\ K-h-203 \\ K-h-203 \\ K-h-203 \\ K-h-203 \\ K-h-201 \\ K-h-203 \\ K-h$		V 6 E 1	PK2PK-f-51
Faulty Data K-F-52 K-f-53 K-f-5n PK2PK-f-53 PK2PK-f-5n 20 MPa. Healthy Data K-h-201 K-h-202 K-h-203 K-h-203 K-h-20n PK2PK-h-202 PK2PK-h-203 PK2PK-h-203 PK2PK-h-201 0.0 20 MPa. Healthy Data K-h-201 K-h-203 K-h-203 K-h-201 PK2PK-h-202 PK2PK-h-203 PK2PK-f-201 20 MPa. K-f-202 MPa. Faulty K-f-201 K-f-203 FK2PK-f-203 PK2PK-f-203 PK2PK-f-203	5 MPa.	K-I-51	PK2PK-f-52
Data K-F-53 PK2PK-f-5n K-f-5n PK2PK-f-5n 20 K-h-201 PK2PK-h-201 MPa. K-h-203 PK2PK-h-203 Healthy K-h-20n PK2PK-h-203 Data K-h-201 PK2PK-h-203 MPa. K-h-20n PK2PK-h-203 MPa. K-h-20n PK2PK-h-201 20 K-f-201 PK2PK-f-201 MPa. K-f-202 PK2PK-f-202 MPa. K-f-203 PK2PK-f-203	Faulty Data	K-I-52	PK2PK-f-53.
K-F-511 PK2PK-f-5n 20 K-h-201 PK2PK-h-201 MPa. K-h-203 PK2PK-h-202 Healthy K-h-203 PK2PK-h-203 Data K-h-201 PK2PK-h-203 20 K-h-203 PK2PK-h-203 MPa. K-h-201 PK2PK-h-203 MPa. K-h-201 PK2PK-f-201 20 K-f-201 PK2PK-f-201 MPa. K-f-203 PK2PK-f-202 MPa. K-f-203 PK2PK-f-203		K-I-55	
20 MPa. Healthy Data K-h-201 K-h-202 K-h-203 PK2PK-h-201 PK2PK-h-202 MPa. Healthy Data FK2PK-h-203 PK2PK-h-203 MPa. K-h-20n PK2PK-h-203 PK2PK-h-203 20 MPa. Faulty K-f-201 K-f-202 PK2PK-f-202 PR2PK-f-203 PK2PK-f-203		K-I-5II	PK2PK-f-5n
20 MPa. Healthy Data K-h-201 K-h-202 K-h-203. K-h-203. PK2PK-h-202 PK2PK-h-203 0.0 PK2PK-h-203 0.0 PK2PK-f-201 20 MPa. Faulty K-f-201 K-f-202 K-f-203 PK2PK-f-202 PK2PK-f-203 PK2PK-f-203	20		PK2PK-h-201
20 MPa. Healthy K-f-201 K-h-202 K-h-203 PK2PK-h-202 PK2PK-h-203 Data K-h-20n PK2PK-h-203 20 MPa. K-f-201 K-f-202 PK2PK-f-201 20 MPa. K-f-201 K-f-202 PK2PK-f-202 MPa. K-f-203 PK2PK-f-203		K b 001	
MPa. Healthy Data K-fi-202 K-h-203. K-h-20n PK2PK-h-203 Fixed PK2PK-h-20n 0.0 PK2PK-f-201 20 MPa. Faulty K-f-201 K-f-202 K-f-203. PK2PK-f-202		K-11-201	PK2PK-h-202
Healthy K-f-203 PK2PK-h-203 Data K-h-20n PK2PK-h-201 K-h-20n PK2PK-f-201 20 K-f-201 MPa. K-f-203 Faulty PK2PK-f-203	MPa.	K-II-202	
Data K-h-20n PK2PK-h-20n 0.0 PK2PK-f-201 20 K-f-201 PK2PK-f-202 MPa. K-f-203 PK2PK-f-203	Healthy Data	K-II-205.	PK2PK-h-203
K-f-201 PK2PK-h-20n 0.0 PK2PK-f-201 20 K-f-201 MPa. K-f-203 Faulty PK2PK-f-203		 K. b. 00-	
0.0 20 K-f-201 PK2PK-f-201 MPa. K-f-203 PK2PK-f-202		K-II-20II	PK2PK-h-20n
20 K-f-201 PK2PK-f-201 MPa. K-f-202 K-f-203 PK2PK-f-202 K-f-203 PK2PK-f-203			0.0
20 K-f-201 PK2PK-f-202 MPa. K-f-203 PK2PK-f-203	20		PK2PK-f-201
20 K-f-201 PK2PK-f-202 MPa. K-f-203 PK2PK-f-203		V. C 001	
MPa. K-f-202 Faulty K-f-203 PK2PK-f-203		K-I-201	PK2PK-f-202
Faulty K-I-203. PK2PK-f-203	MPa.	K-f-202	
	Faulty Data	K-f-203	PK2PK-f-203
Data Transa		 K. 6.00-	
K-I-20n PK2PK-f-20n		K-t-20n	PK2PK-f-20n
0.0			0.0

TABLA 1 Estructura de la matriz de entrada de la a MATLAB función de K-media

Autoría Propia

La matriz resultante es una matriz de dos columnas que contiene la información de las dos características para los cuatro escenarios que tenemos. K se utiliza para representar los valores de curtosis y PK2PK para representar los valores de pico a pico. "K" y "PK2PK" van seguidos de una "h" o una "f" que significa "en buen estado" y "defectuoso" respectivamente. Al final, es un "5" o un "20" para representar los valores correspondientes a las presiones de 5 MPa y 20 MPa respectivamente. Los subíndices representan la posición o el número del ciclo. Como se mencionó anteriormente, no se consideraron los ciclos que presentaron el error de cambio cero.

Resultados de la implementación: Si bien se obtuvo un buen resultado con los datos de la curtosis de la corriente y los valores pico a pico de la aceleración (figura 26), hubo algunos puntos que fueron identificados erróneamente por el algoritmo. Una de las razones para que esto suceda es la diferencia del rango en los valores de las dos características. Por ejemplo, podemos apreciar que los valores pico a pico están alrededor de 5 y 115, mientras que los valores de la curtosis están alrededor de 2 y 12. Esto es un problema ya que el algoritmo usa la distancia de cada punto al centroide para clasificar los puntos a los diferentes grupos.

Diego Yánez, et al. Uso de algoritmo K-media para detectar estado de funcionamiento de electroválv...



Este problema se puede resolver normalizando los valores de las características. Al normalizar los valores, podemos volver a escalarlos en un rango entre 0 y 1. Para realizar esta acción se utiliza la ecuación 8:



Autoría Propia

También vale la pena agregar el análisis para la fase C. El proceso de extracción de características también mostró que las mismas características utilizadas para la fase A (curtosis y pico a pico) son útiles para agrupar los datos pertenecientes a la fase C. Los gráficos 26, 27 y 28 muestran las características de la fase C, por lo que la función k-media puede distinguir entre saludable y defectuoso.



Se obtienen resultados bastante parecidos una vez normalizada la data. De esta manera se puede comprar los datos de las fases A y C, con una distinción entre saludable y defectuoso.

Agradecimientos

Agradezco de sobremanera a mi familia, dado su amplio sacrificio en el apoyo a mi formación tanto profesional y como persona. (Yánez, D)

Referencias Bibliográficas

- Alqadi, Z., & ... (2019). Adaptation of matlab K-means clustering Al-Azzeh, J., Rasras, R., function to create Color Image Features. ... Journal of Research ..., May, 10 -18. https://www.researchgate.net/profile/Ziad_Alqadi/publication/333402203_Adaptation_of_matlab_Kmeans_clustering_function_to_create_Color_Image_Features/links/5cebe595299bf14d95bdc749/ Adaptation-of-matlab-K-means-clustering-function-to-create-Color-Image-Fe
- García, C., & Gómez, I. (2006). Algoritmos de aprendizaje: knn & kmeans. Universidad Carlos III de Madrid. http://www.it.uc3m.es/jvillena/irc/practicas/08-09/06.pdf
- Har-peled, S. (2004). On Coresets for k -Means and k -Median Clustering. 291–300.
- Hernández Sampieri, R., & Mendoza Torres, C. P. (2018). Metodología de la investigación: las tres rutas cuantitativa, cualitativa y mixta. In Mc Graw Hill (Vol. 1, Issue Mexico).
- Milone, D., Stegmayer, G., Gerard, M., Kamenetzky, L., López, M., & Carrari, F. (2009). Métodos de agrupamiento no supervisado para la integración de datos genómicos y metabólicos de múltiples líneas de introgresión. Inteligencia Artificial, 13(44), 56–66. https://doi.org/10.4114/ia.v13i44.1046
- Parker Corporation, H. (2007). Bulletin HY14-2536-M1/US, 600, 4/07, PHD. Parker Hannifin Corporation Hydraulic Valvule Division, 1.
- Russell, S., & Norvig, P. (2007). Inteligencia Artificial. Inteligencia Artificial, 2(6), 1–33. https:// static0planetadelibroscom.cdnstatics.com/libros_contenido_extra/40/39308_Inteligencia_artificial.pdf %0Ahttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Inteligencia+Artificial:+un +enfoque+moderno#0
- Song, T. T., & Huang, N. (2005). A novel zero-voltage and zero-current-switching full-bridge PWM converter. IEEE Transactions on Power Electronics, 20(2), 286–291. https://doi.org/10.1109/TPEL.2004.843016