

Diagnóstico de fallas mediante una LSTM y una red elástica

Márquez-Vera, M.A.^{a, *}, López-Ortega, O.^b, Ramos-Velasco, L.E.^c, Ortega-Mendoza, R.M.^d,
Fernández-Neri, B.J.^a, Zúñiga-Peña, N.S.^{a,b}

^aUniversidad Politécnica de Pachuca, C. Pachuca-Cd. Sahagún Km 20, C.P. 43830 Zempoala, Hgo., México, Tel/Fax. (+52)771 5477510.

^bUniversidad Autónoma del Estado de Hidalgo, C. Pachuca-Tulancingo Km 4.5, C.P. 42090 Mineral de la Reforma, Hgo., México.

^cUniversidad Politécnica Metropolitana de Hidalgo, Boulevard Acceso a Tolcayuca 1009, Ex-Hacienda de San Javier, C.P. 43860, Tolcayuca, Hgo., México.

^dUniversidad Politécnica de Tulancingo, Calle Ingenierías #100, Huapalcalco, C.P. 43629 Tulancingo de Bravo, Hgo., México

To cite this article: Márquez-Vera, M.A., López-Ortega, O., Ramos-Velasco, L.E., Ortega-Mendoza, R.M., Fernández-Neri, B.J., Zúñiga-Peña, N.S. 2021. Fault diagnosis in industrial process by using LSTM and an elastic net. Revista Iberoamericana de Automática e Informática Industrial 18, 164-175. <https://doi.org/10.4995/riai.2020.13611>

Resumen

El diagnóstico de fallas es importante en los procesos industriales, ya que permite determinar si es necesario detener el proceso en operación y/o proponer un plan de mantenimiento. En el presente trabajo se comparan dos estrategias para diagnosticar fallas. La primera realiza un preprocesamiento de datos usando el análisis de componentes independientes para reducir la dimensión de los datos, posteriormente, se emplea la transformada wavelet para resaltar las señales de falla, con esta información se alimenta una red neuronal artificial. Por su parte, la segunda estrategia, principal contribución de este trabajo, usa una memoria de corto y largo plazo. Esta memoria es alimentada por las variables más significativas seleccionadas mediante una red elástica para usar tanto la norma L_1 como la L_2 . Como ejemplo de aplicación se utilizó el proceso químico Tennessee Eastman, un proceso ampliamente usado en el diagnóstico de fallas. El aislamiento de fallas mostró mejores resultados con respecto a los reportados en la literatura.

Palabras clave: Diagnóstico de fallas, Transformada Wavelet, Redes neuronales recurrentes, Análisis de componentes independientes, Red elástica.

Fault diagnosis in industrial process by using LSTM and an elastic net

Abstract

Fault diagnosis is important for industrial processes because it permits to determine the necessity of emergency stops in a process and/or to propose a maintenance plan. Two strategies for fault diagnosis are compared in this work. On the one hand, the data are preprocessed using the independent components analysis for dimension reduction, then the wavelet transform is used in order to highlight the faulty signals, with this information an artificial neural network was fed. On the other hand, the second strategy, the main contribution of this work, is the implementation of a long short term memory. This memory is fed with the most representative variables selected by an elastic net to use both, the L_1 and L_2 norms. These strategies are applied in the Tennessee Eastman process, a benchmark widely used for fault diagnosis. The fault isolation had better results than those reported in the literature.

Keywords: Fault diagnosis, Wavelet transform, Recurrent neural networks, Independent component analysis, Elastic net.

1. Introducción

El diagnóstico de fallas, comúnmente conocido como FDI por sus siglas en inglés (Zhang et al., 2002), consta de dos partes: (i) detección de fallas y (ii) aislamiento de fallas. La detección de fallas consiste en notar que un mal funcionamiento

ha ocurrido (Salahshoor et al., 2011). Por su parte, el aislamiento de fallas, también conocido como clasificación de fallas (Sobhani-Tehrani and Khorasani, 2009) o localización de fallas (Barakat et al., 2014), consiste en reconocer cuál fue la falla que ocurrió entre varias fallas posibles (Jing et al., 2014).

*Autor para correspondencia: marquez@upp.edu.mx

Cuando se cuenta con un modelo aproximado del proceso a analizar, es posible calcular la diferencia entre la salida entregada por el modelo aproximado y la salida real del proceso, tal diferencia se conoce como residuo. Si la diferencia sobrepasa un determinado umbral, se establece que se ha detectado una falla (Kandula, 2011). Comúnmente, cuando varias señales se alejan de los valores esperados, se realiza una clasificación de residuos para aislar las fallas (Linker et al., 2002; Methnani et al., 2013).

En procesos complejos donde no se puede proponer un modelo aproximado, es posible realizar el FDI empleando métodos basados en datos. En las siguientes subsecciones se describen los métodos comúnmente aplicados para el FDI.

1.1. Métodos estadísticos en FDI

Una alternativa es usar técnicas estadísticas, la más común es el análisis de componentes principales (PCA, por sus siglas en inglés). Este análisis consiste en hallar las direcciones de mayor varianza entre las variables de entrada con el fin de quitar aquellas variables linealmente dependientes de las demás y que presentan mayor covarianza (Lau et al., 2013; Gao y Hou, 2016). Otra estrategia es el análisis de componentes independientes (ICA, por sus siglas en inglés), el cual puede tratar con datos que no presentan una distribución gaussiana, pues busca las direcciones de mayor independencia (Salahschoor y Kiasi, 2008; Zhang y Zhao, 2017).

Comúnmente, los métodos de diagnóstico de fallas son evaluados usando casos de estudio como el proceso químico Tennessee Eastman (TE) (ver Sección 2). Este proceso ha sido empleado ampliamente para probar técnicas de control tolerante a fallas y diagnóstico de fallas (Chen et al., 2014). Particularmente, Yin et al. (2012) realizaron una comparación interesante entre el PCA y el ICA aplicados al FDI del proceso TE. En su trabajo, también comparan algunas de sus variantes como el análisis discriminante de Fisher (FDA) y mínimos cuadrados parciales (PLS). En su Tabla 7, se muestra que las fallas denominadas IDV3, IDV9 e IDV15 fueron las más difíciles de detectar, y son nombradas fallas incipientes. Resultados similares fueron mostrados por (Salahschoor y Kiasi, 2008).

Generalmente, después de procesar los datos por PCA o ICA, se calculan las estadísticas T^2 o el error de predicción cuadrático (SPE , por sus siglas en inglés) (Zhang et al., 2012). Posteriormente, se fija un umbral, si la estadística sobrepasa el umbral, se detecta una falla. Sin embargo, debido al ruido en las mediciones, es posible que existan falsas alarmas cuando los umbrales fijados son sobrepasados accidentalmente. Para enfrentar este problema, Alkaya y Eker (2011) propusieron un umbral dinámico obteniendo una detección de fallas dinámica llamada DPCA o DICA (Rato y Reis, 2013).

1.2. Uso de wavelets en FDI

Una manera de enriquecer la información a procesar consiste en resaltar eventos en las señales mediante la transformada wavelet. Este tipo de transformada trata con señales no estacionarias y proporciona información en el dominio del tiempo así como en el dominio de la frecuencia. Debido al análisis frecuencial de las wavelets, la mayoría de sus aplicaciones se asocia con sistemas que muestran un comportamiento periódico en

las mediciones, tal como sucede en los motores eléctricos, generadores eléctricos o en cajas de engranes (Wu et al., 2009; Khakipour et al., 2017; Lou y Loparo, 2004). Una revisión sobre el uso de wavelets para diagnóstico de fallas en maquinaria rotatoria se encuentra en el trabajo realizado por Yan et al. (2014).

También se han realizado algunas aplicaciones de wavelets para tratar otro tipo de sistemas. Por ejemplo, en el trabajo presentado por Adewole et al. (2016) se emplean wavelets discretas para resaltar determinados detalles en las señales medidas para localizar fallas en redes de distribución de energía. De manera similar, Jung et al. (2007) localizaron fallas en cables subterráneos procesando la información con wavelets y redes neuronales. Por su parte, Nguyen et al. (2017) emplearon wavelets y redes neuronales para reconocer fallas en una subestación eléctrica en Vietnam.

Por otro lado, Wu et al. (2011) usaron wavelets difusas para diagnosticar fallas en una línea de ensamble de automóviles; mediante el uso de algoritmos genéticos, los autores mejoraron su propuesta requiriendo menos muestras y con una mayor tasa de detección de fallas. Wu et al. (2016) aplicaron el ICA a la señal de un electromiograma, posteriormente, usaron las wavelets como filtro con el fin de diagnosticar enfermedades respiratorias. Referente al proceso TE, Salahschoor y Kiasi (2008) emplearon la transformada wavelet rápida para quitar el ruido en las variables medidas para después emplear ICA en la supervisión del proceso.

1.3. Selección de variables

Actualmente, existen métodos que aplican selección de variables en el diagnóstico de fallas. La idea es conocer anticipadamente al conjunto de variables más representativas para realizar el FDI. En consecuencia, es posible emplear menos instrumentación en el proceso y procesar menos información. Cabe señalar que, reducir la dimensión de datos usando PCA o ICA es un concepto diferente de seleccionar las variables más representativas. En el primer caso se trata de una aproximación lineal del sistema, mientras que la selección de variables se enfoca en eliminar la redundancia entre ellas.

Una técnica conocida para seleccionar las variables es la conocida como información mutua. En este contexto, Zhang y Zhao (2017) propusieron utilizar 10 variables del proceso TE para aplicar una red de creencia profunda con el objetivo de diagnosticar fallas. Otra alternativa es usar regresión múltiple, si se agrega una penalización usando la norma L_1 se obtiene un modelo llamado contracción de mínimos absolutos y operadores de selección (LASSO, por sus siglas en inglés). Un algoritmo muy similar al LASSO pero más eficiente es la regresión por ángulos mínimos (LARS) (Hastie et al., 2009). Si a este último se agrega una penalización con la norma L_2 se obtiene el modelo de red elástica (EN, por sus siglas en inglés) (Kuang et al., 2015).

Una vez conocidas las variables más representativas, el FDI es obtenido mediante algoritmos de clasificación como Redes Neuronales Artificiales (ANN, por sus siglas en inglés). Un trabajo completo sobre el uso de la ANN en el diagnóstico de fallas se presenta en (Patan, 2008). Recientemente, también ha surgido el interés de aplicar aprendizaje profundo.

1.4. Aprendizaje profundo en FDI

Si bien las ANN son empleadas para el aislamiento de fallas (Zhao, 2018; Oliveira et al., 2017), también es posible realizar una red neuronal profunda para extraer características con diferentes niveles de abstracción. Un problema que se genera en este tipo de redes es el desvanecimiento del gradiente, es decir, que mientras más capas tenga una red neuronal, es cada vez más difícil actualizar los pesos de las capas más cercanas a la entrada usando métodos de entrenamiento como la retro-propagación del error (Zhang et al., 2018).

Para poder entrenar una ANN profunda se requiere un cambio de paradigma. Una alternativa es emplear una máquina restringida de Boltzmann (RBM, por sus siglas en inglés). Las máquinas de Boltzmann tienen varios nodos completamente conectados unos con otros. Si se restringen las conexiones entre nodos que pertenecen a una misma capa se obtiene una RBM, la cual tiene dos capas: la capa visible y la oculta (Geng et al., 2018). En este paradigma, la idea es que la capa oculta reproduzca la información proveniente de la capa visible usando una función de la densidad de probabilidad para ajustar los pesos sinápticos entre capas.

La técnica de aprendizaje profundo que emplea varias RBMs apiladas es la red de creencia profunda (DBN, por sus siglas en inglés). En esta red, se suele agregar una última capa llamada capa de características. Para entrenar una DBN, se ajustan las RBM tomando dos capas adyacentes a la vez. Este proceso es conocido como pre-entrenamiento. Después se ajustan las últimas dos capas como una ANN clásica. Finalmente, se entrena toda la red empleando el gradiente del error, esta fase se conoce como sintonización fina. Estas redes han sido aplicadas en FDI. Por ejemplo, Zhang y Zhao (2017) usaron una DBN para el diagnóstico de fallas en el proceso TE. En su trabajo propusieron una función gaussiana para poder usar valores continuos en la función de energía usada en el pre-entrenamiento.

Otro tipo de estructura usada en aprendizaje profundo es el auto-sintonizador para quitar ruido (SDA, por sus siglas en inglés), el cual de manera similar a la RBM, se entrena usando dos capas a la vez. En cada pareja se agrega ruido a las variables de entrada (codificación) y la otra capa debe recobrar la señal original (decodificador). El entrenamiento de este tipo de capas (capas sabias) forma el pre-entrenamiento, siendo hasta este punto, un sistema no supervisado (Lv et al., 2016). Finalmente, el aislamiento de fallas se puede lograr con la sintonización fina. El uso de SDAs en el diagnóstico de fallas ha sido mostrado por Lv et al. (2017). Ellos emplearon una serie de tiempo para determinar las correlaciones entre las fallas, teniendo la posibilidad de abstraer características de alto nivel para detectar fallas incipientes en el proceso TE. Mientras, Lv et al. (2018) propusieron mejorar la detección de fallas en el proceso TE añadiendo una restricción de dispersión en los SDAs.

Las Redes Neuronales Convolucionales (CNN, por sus siglas en inglés) son otro esquema de aprendizaje profundo ampliamente utilizado en el procesamiento de imágenes y video. La función de este tipo de redes es extraer características de los datos procesados, así como clasificarlos. Estas redes cuentan con capas de convolución que generan como salida un mapa de características obtenido mediante un filtro. Todas las neuronas en un mapa de salida comparten el mismo filtro. Por su parte, las capas de agrupación o sub-muestreo, obtienen como salida

un mapa de características con menos datos agrupando características locales similares (Smirnov et al., 2014). De manera similar a los esquemas mencionados, estas redes contienen una capa final de neuronas completamente conectadas para clasificar información; esta capa es entrenada con la sintonización fina. En el contexto del diagnóstico de fallas, Wu et al. (2018) aplicaron una CNN en el proceso TE para extraer características en el dominio del tiempo y el espacio, sus resultados muestran una tasa de detección de fallas en promedio de 88.2 % para las 20 fallas que se pueden simular en el proceso. Cabe señalar que su propuesta es válida para procesos industriales que cuenten con un historial de mediciones.

Recientemente, las redes recurrentes, otro modelo de aprendizaje profundo, han comenzado a utilizarse para FDI, sobresaliendo las redes de memoria a corto y largo plazo (LSTM, por sus siglas en inglés). Las LSTM están constituidas por celdas de memoria, cuyo estado es regulado por las compuertas de entrada, de salida, de candidatura y de olvido. Cada compuerta está compuesta por una red neuronal completamente conectada (Xiao et al., 2019), y en la etapa de entrenamiento se trunca el gradiente cuando éste deja de ser relevante (Hochreiter y Schmidhuber, 1997). Las aplicaciones de estas redes han sido diversas. Por ejemplo, Lei et al. (2019) aplicaron una LSTM para diagnosticar fallas en una turbina generadora de electricidad. La idea fue tratar con series de tiempo, y de esta manera, determinar la dependencia entre términos largos para clasificar fallas.

En el presente trabajo se propone el uso de una LSTM para la detección y aislamiento de fallas. Con el fin de acelerar el aprendizaje de la LSTM. Otra alternativa sería usar la transferencia de aprendizaje para ajustar la red por capas. La selección de variables más representativas se lleva a cabo mediante una red elástica, y para fines de comparación se presenta también una ANN entrenada con datos procesados con la transformada wavelet. Estos datos son obtenidos de las componentes independientes seleccionadas con ICA.

El resto del artículo está organizado de la siguiente manera. En la Sección 2 se describe brevemente el proceso TE como caso de aplicación para evaluar los métodos propuestos. La Sección 3 presenta la metodología propuesta para la detección y el aislamiento de fallas, primero se muestra la ANN diseñada y posteriormente, la arquitectura del modelo LSTM propuesto. Los resultados de ambos enfoques se muestran en la Sección 4. Finalmente, la Sección 5 resume las conclusiones.

2. Proceso Tennessee Eastman: caso de aplicación

El proceso TE ha sido ampliamente utilizado como marco de referencia para comparar diferentes algoritmos de FDI (Yin et al., 2012; Lau et al., 2013; Chen et al., 2014; Boldt et al., 2014). Este proceso fue presentado por Downs y Vogel (1993). La idea es ofrecer un modelo que represente un proceso industrial real sobre el cual proponer estrategias para el monitoreo de procesos, detección y diagnóstico de fallas, así como técnicas de control tolerante a fallas.

Este proceso cuenta con cinco partes principales: un reactor de dos fases donde ocurre una reacción exotérmica, un separador de líquido-vapor, un condensador del producto, un compresor y un depurador de producto. Tiene ocho componentes

llamados A, B, ..., H (Yin et al., 2012); los componentes A, C, D y E son los reactantes, y los dos productos obtenidos son los líquidos G y H. Mientras se produce la reacción se obtiene el componente F; el inerte B se encuentra en los flujos de alimentación de A, C y D. El proceso TE cuenta con 41 variables que pueden ser medidas llamadas x_1, x_2, \dots, x_{41} . Las primeras 22 tienen un periodo de muestreo de 3 minutos, las demás son variables compuestas con periodos de muestreo de 6 a 15 minutos. Además, tiene 11 variables de entrada, de manera que no es fácil obtener un modelo analítico del proceso completo.

En el proceso TE se pueden simular 20 fallas denominadas IDV1, IDV2, ..., IDV20, las cuales están descritas detalladamente en el trabajo de Downs y Vogel (1993). Un esquema del proceso puede ser apreciado en la Figura 1 de (Jing et al., 2014), donde se comparan los resultados de emplear PCA y SVM en el diagnóstico de fallas. Una revisión al proceso TE se presenta en (Bathelt et al., 2015) estudiando las inconsistencias de emplear diferentes métodos de integración para simular el proceso¹. Naturalmente, algunas fallas son más difíciles de detectar debido a su efecto imperceptible en las variables de salida; a estas fallas se les denomina *incipientes*. Por ejemplo, las fallas IDV3, IDV9 e IDV20 fueron difíciles de detectar en el trabajo de Boldt et al. (2014) mediante el uso de métodos estadísticos. Resultados similares fueron obtenidos por Salahshoor y Kiasi (2008), quienes emplearon ICA con filtros wavelet en el proceso TE.

En el presente trabajo, para detectar y aislar fallas en el proceso TE se realizaron 63 simulaciones con 800 muestras en cada una, de las cuales 200 eran datos sin falla. En las 600 restantes se aplicó una falla en particular. En el modelo se pueden aplicar 20 fallas distintas, de modo que 42 simulaciones fueron usadas para entrenamiento, dos de ellas con 800 muestras sin aplicar ninguna falla. Para validar los resultados, con los métodos a continuación descritos, se emplearon 21 simulaciones más de cada tipo de falla, así como una simulación sin fallas.

3. Metodología propuesta para la detección y el diagnóstico de fallas

En esta sección presentamos dos enfoques para el diagnóstico de fallas. Particularmente, la subsección 3.1 muestra la ANN propuesta, donde los datos medidos son procesados inicialmente mediante ICA. Por su parte, la subsección 3.2 describe el modelo LSTM propuesto.

3.1. Enfoque basado en una ANN

En la ANN propuesta, los datos medidos son procesados inicialmente mediante ICA. Esta técnica permite quitar la dependencia lineal entre variables, además de reducir la dimensión de los datos agilizando el aprendizaje de la red. Adicionalmente, para resaltar los eventos de falla en los datos medidos se emplea la transformada wavelet.

3.1.1. Análisis de componentes independientes

El análisis de componentes independientes (ICA) busca una representación lineal de variables no gaussianas (Yin et al.,

2012), de manera que las componentes encontradas sean independientes estadísticamente. La representación formada con las componentes encontradas tendrá la estructura esencial de los datos originales. Usando ICA es posible separar señales obtenidas de la mezcla de varias fuentes (Lin y Zhang, 2005). La independencia entre variables significa que no se puede inferir información sobre las demás componentes empleando una componente en especial. Este concepto suele definirse mediante funciones de densidad de probabilidad, i.e. si denotamos como $p(x_1, x_2)$ a la función de probabilidad conjunta de x_1 y x_2 , y sea $p_1(x_1)$ la función de densidad de probabilidad marginal cuando x_1 es considerada sola, siendo obtenida con (1)

$$p_1(x_1) = \int p(x_1, x_2) dy_2, \quad (1)$$

y de forma similar para el caso de x_2 . Entonces se dice que x_1 y x_2 son independientes, si y sólo si su función de densidad de probabilidad conjunta $p(x_1, x_2)$ es factorizable, i.e.

$$p(x_1, x_2) = p_1(x_1)p_2(x_2). \quad (2)$$

El ICA parte de una representación de variables latente, para la cual se normalizan las mediciones tomadas con el objetivo de obtener una media cero. Tales mediciones son las columnas de la matriz $X \in \mathbb{R}^{N \times m}$, donde N el número de muestras y m la cantidad de variables que se miden. Posteriormente, se obtiene la descomposición en valores singulares (SVD, por sus siglas en inglés), $X = \frac{1}{\sqrt{N+1}}(X^T X) = UDV^T$, siendo $D \in \mathbb{R}^{N \times m}$ una matriz cuya diagonal contiene los valores propios de X . A su vez, $U \in \mathbb{R}^{N \times N}$ está formada por los vectores propios de X .

Mediante las matrices $S = \sqrt{N}U$ y $A^T = DV^T / \sqrt{N}$ se obtiene que $X = SA^T$. Por lo tanto, las columnas de X son una combinación lineal de las columnas de S que también tienen media cero. Describiendo de manera extendida se obtiene el conjunto (3) (Hastie et al., 2009)

$$\begin{aligned} X_1 &= a_{11}S_1 + a_{12}S_2 + \dots + a_{1m}S_m, \\ X_2 &= a_{21}S_1 + a_{22}S_2 + \dots + a_{2m}S_m, \\ &\vdots \\ X_m &= a_{m1}S_1 + a_{m2}S_2 + \dots + a_{mm}S_m, \end{aligned} \quad (3)$$

el cual se puede reescribir como $x = As$. Más aún, si se emplea un valor $d < m$, se obtiene una descomposición trunca para aproximar X de una manera más simple. De esta manera, s debe ser linealmente independiente. Para ello, todos los momentos cruzados son tomados en cuenta y no sólo los de segundo grado. De esta manera, se encuentra una la solución para A .

Una forma de interpretar el ICA es mediante el problema del Cóctel, donde se supone que existen varias fuentes de señales captadas por diversos sensores S . Las señales obtenidas de los sensores resultan ser una combinación lineal de las señales provenientes de todas las fuentes X (Yin et al., 2012). El problema radica en recuperar las señales originales a partir de los sensores. Por esta razón, la matriz A es llamada matriz de mezclado (Hastie et al., 2009). Si los datos de entrada x son blanqueados mediante SVD para tener covarianza unitaria, i.e.

¹El modelo puede ser descargado de manera gratuita desde la página <http://depts.washington.edu/control/LARRY/TE/download.html>

$cov(x) = I$, se tiene que A es ortogonal porque s también tendrá covarianza unitaria.

Para determinar el ICA se suele emplear la entropía de información de Claude Shannon H (Muñoz-Cobo et al., 2017); la cual, al estar definida para alguna variable aleatoria Y con la densidad $p(y)$, está dada por (4)

$$H(Y) = - \int p(y) \log p(y) dy, \quad (4)$$

además, de la teoría de la información, se sabe que entre diversas variables con la misma varianza, las que tienen mayor entropía son las variables gaussianas (Goodfellow et al., 2016).

Una manera de medir la dependencia entre variables es usando la información mutua (MI, por sus siglas en inglés), definida mediante (5)

$$I(Y) = \sum_{i=1}^p H(Y_i) - H(Y) - \log |\det A|. \quad (5)$$

Zhang y Zhao (2017) emplearon la MI para seleccionar las variables más representativas en el proceso TE, las cuales se obtuvieron al aplicar cada una de las 20 fallas simuladas para realizar el diagnóstico de fallas. La MI es llamada también la distancia Kullback-Leibler entre la densidad $p(y)$ y su versión independiente $\prod_{i=1}^p p_i(Y_i)$, donde p_i es la densidad marginal de Y_i .

Para encontrar A es posible utilizar la entropía negativa $J(Y_i)$ (Hyvärinen y Oja, 2000) definida por (6)

$$J(Y_i) = H(Z_i) - H(Y_i), \quad (6)$$

donde Z_i es una variable aleatoria gaussiana con la misma varianza que Y_i . Para implementar el ICA se puede realizar la aproximación (7)

$$J(Y_i) \approx [E(G(Y_i)) - E(G(Z_i))]^2, \quad (7)$$

donde E representa la esperanza matemática y G representa alguna función no cuadrática propuesta como (8) o (9) en el algoritmo llamado *FastICA* (Lin y Zhang, 2005).

$$G_1(u) = \tanh(a_1 u), \quad (8)$$

$$G_2(u) = u \exp(-u^2/2), \quad (9)$$

Para el caso donde $1 \leq a_1 \leq 2$, la aproximación (7) evita utilizar la curtosis (Rockinger y Jondeau, 2002) para calcular la entropía negativa.

Una manera de implementar el ICA fue presentada por Yin et al. (2012) como una serie de pasos, los cuales se resumen a continuación:

1. Se adquieren N muestras $X^T = [x_1, x_2, \dots, x_N]$, se centran para tener media cero, y se calcula

$$\chi = QX^T, \text{ siendo } Q = D^{-1/2}V,$$

2. Se iteran los siguientes cálculos m veces

$$J(Y) \approx [E(G(Y)) - E(G(Z))]^2,$$

$$b_k = \arg \max_{y \in Y} J(y), \quad y = b_k^T \chi,$$

después, se forma la matriz B con todas las b_k , para obtener la matriz W , la cual quita la mezcla de las mediciones mediante $W = B^T Q$.

3. Se determina el número de componentes independientes d para tener las matrices asociadas que quitan el mezclado $W_d \in \mathbb{R}^{d \times m}$ y la matriz residual $W_r \in \mathbb{R}^{m-d \times m}$. En (Lee et al., 2004) se muestra una manera de proponer la cantidad de componentes independientes (IC, por sus siglas en inglés). En el presente trabajo propusimos $d = 5$ con base en las magnitudes de los valores propios mostrados en la Figura 1. Usando las matrices obtenidas, se emplean las siguientes estadísticas (Yin et al., 2012):

- Sumatoria de las componentes dominantes

$$I_d^2 = x^T W_d^T W_d x,$$

- Sumatoria de las componentes residuales

$$I_r^2 = x^T W_r^T W_r x,$$

Propuesta por Lee et al. (2004) y adoptada en varios trabajos similares (Žvokelj et al., 2016).

- Error cuadrático de predicción

$$SPE = e^T e.$$

4. Se determinan los umbrales para detección con base en la estadística usada J_{I_d} , J_{I_r} y J_{SPE} .
5. Si se cumplen las siguientes desigualdades, se asume que no se ha presentado ninguna falla, en caso contrario ha ocurrido una: $SPE \leq J_{SPE}$ y $I_d^2 \leq J_{I_d}$ y $I_r^2 \leq J_{I_r}$.

Empleando el FastICA con las 41 variables medibles del proceso TE, se seleccionaron cinco componentes. Lo anterior debido a la magnitud de las primeras cinco componentes comparadas con el resto. Como se aprecia en la Figura 1, los valores propios de las primeras cinco componentes retenían el 78 % de la información.

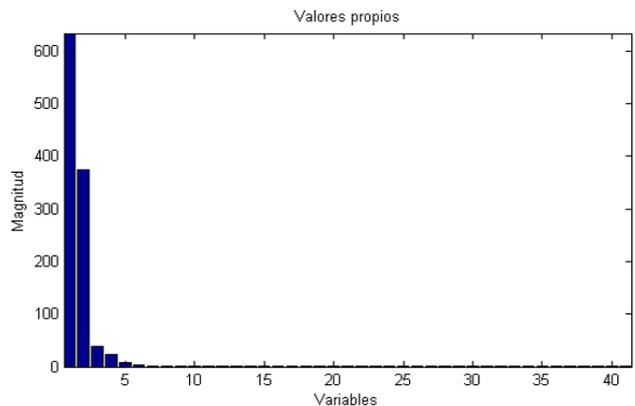


Figura 1: Magnitudes de las componentes independientes.

Una vez determinadas las componentes principales, los transitorios en las señales debidas a las fallas pueden ser resaltados usando la transformada wavelet, la cual es descrita en la siguiente Sección.

3.1.2. Procesamiento de las componentes independientes con wavelets

Empleando el ICA se obtienen las componentes independientes que llevan la mayor parte de información, al ser distintas entre sí, la huella dejada por aquellas fallas que no sean detectables en una variable, pueden ser buscadas en otra. Para resaltar la información obtenida empleando ICA se propone realizar un procesamiento con la transformada wavelet. Este procesamiento se fundamenta en la idea de Liang et al. (2019) que hace referencia a la transformada wavelet como un microscopio matemático para el análisis de señales. Dado que las señales de falla no son estacionarias, la transformada wavelet es una buena opción, pues mapea la señal en el dominio del tiempo y de la frecuencia. El mapeo se realiza mediante un corrimiento y una dilatación de una wavelet llamada *madre* (Ramos-Velasco et al., 2013). La transformada wavelet continua de una función $f(t)$ está dada por (10)

$$\langle f, \psi_{a,b} \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt, \quad (10)$$

donde $\psi(\cdot)$ representa a la wavelet madre, a es la escala y b es el corrimiento. Las condiciones para que $\psi(\cdot)$ sea una base wavelet son las siguientes:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0,$$

$$\int_{-\infty}^{\infty} \psi^2(t) dt = 1.$$

Las wavelets pueden verse como filtros pasa altas resaltando los transitorios de una señal (Li et al., 2014). Su complemento son las funciones de escala que pueden ser vistas como filtros pasa bajas, siendo estas funciones ortogonales a las wavelets.

Ojeda-González et al. (2014) mencionan que no existe una selección perfecta de la transformada wavelet a utilizar. En este sentido, existen varios trabajos que comparan los resultados del uso de diferentes transformadas wavelet. Por ejemplo, en (Rodríguez et al., 2019) se clasifican señales EEG para detectar personas con alcoholismo empleando las transformadas *wavelet Symlet*, *Coiflet*, *Biorthogonal* y *Daubechies*; los resultados fueron muy similares. Sharif and Khare (2014) clasificaron imágenes empleando las wavelets *Haar* y *Daubechies*. En (Kumar y Bansal, 2019) se comparan los resultados de emplear las transformadas wavelet *Haar*, *Daubechies*, *Symlet* y *Biorthogonal* para mitigar los problemas que afectan la calidad en el suministro de energía eléctrica; el mejor resultado se obtuvo con la wavelet *Haar*. En (Dixit y Majumdar, 2013) se comparan los resultados de reducir ruido en señales usando las wavelets *Coiflet* y *Daubechies*.

El uso de la transformada wavelet permite, principalmente, detectar variaciones en una señal que representen una falla, incluso cuando el ruido tenga una magnitud mayor que las variaciones analizadas. Lo anterior es útil para detectar fallas incipientes. Además del tipo de wavelet, existen diferentes grados de aproximación y de descomposición. En el trabajo de Rafiee et al. (2010) se realizó el diagnóstico de fallas en rodamientos y se seleccionó el tipo de wavelet así como el orden de la misma para extraer características en las señales que revelaran la presencia de una falla; la *Daubechies44* se posicionó como la mejor opción.

La implementación de métodos que emplean wavelets e ICA han sido desarrollados en el diagnóstico de fallas en engranajes y rodamientos. Por ejemplo, Lin y Zhang (2005) utilizaron las wavelets para generar más señales con el fin de detectar fallas empleando un solo sensor. Por otra parte, Han et al. (2015) obtuvieron las componentes principales y posteriormente las procesaron mediante filtros wavelet.

Una manera de seleccionar el tipo y la escala de wavelet para FDI es mediante la inspección de diferentes escalogramas obtenidos al aplicar diferentes wavelets. Tomando en cuenta la forma de las posibles wavelet madre, se aprecia que la llamada *sombrero mexicano* no presenta tantas crestas que dificultan la clasificación de las señales procesadas como algunas otras (*Morlet* o *Daubechies40*). En la Figura 2 se muestra el resultado de aplicar la transformada wavelet *sombrero mexicano* a la primer componente independiente obtenida mediante ICA cuando se aplicaron 13 fallas distintas en diferentes instantes. La wavelet madre está definida por (11),

$$\psi(x) = \frac{2}{\sqrt{3}} \pi^{-\frac{1}{4}} (1-x^2) e^{-\frac{x^2}{2}}. \quad (11)$$

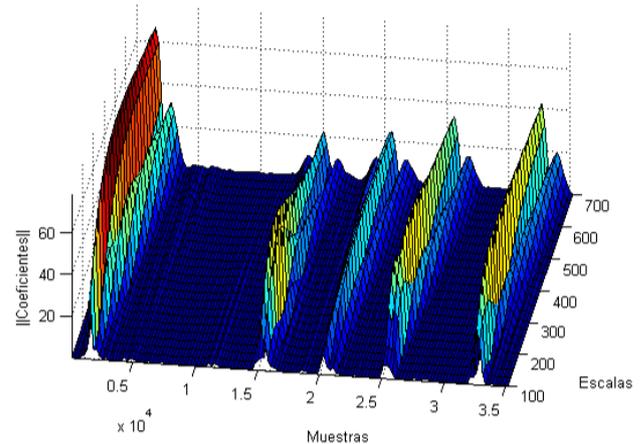


Figura 2: Transformada wavelet *sombrero mexicano* de la primer componente independiente (Figura 1)

En el presente trabajo se emplea la wavelet tipo *sombrero mexicano*, ya que como se mencionó en (Lin y Zhang, 2005), las wavelets no ortogonales funcionan como filtros, y al aplicar ICA a las señales filtradas, no habrá correlación en cualquiera de los momentos de alto orden para señales independientes. Además, la contribución de una misma señal en diferentes mediciones será diferente de cero si los filtros empleados son no ortogonales.

3.1.3. Red neuronal artificial para FDI

El diagnóstico de fallas ha sido considerado como un problema de discriminación de datos para buscar anomalías en las señales medidas. Por tal motivo, (Lv et al., 2016) tratan el FDI como un reconocimiento de patrones. Las ANN han sido empleadas para procesar la información en busca de algún patrón en particular (Maglaveras et al., 1998; Yao et al., 2019).

Para tener diferentes niveles de abstracción de información en busca de algún patrón, las ANN con varias capas son una alternativa, ya que al reducir el número de neuronas capa tras capa, diferentes grados de descomposición son obtenidos (Hoang y Kang, 2019). De esta manera, se obtiene una representación jerárquica de la información análoga a los diferentes niveles de descomposición de una señal mediante la transformada wavelet (Zhao et al., 2019).

El aislamiento de fallas fue realizado empleando una ANN hacia adelante de cinco entradas y cuatro salidas. Las señales obtenidas de la convolución con las wavelets se normalizaron para emplear la tangente hiperbólica como función de activación en las 20 neuronas de la capa oculta de la red. Cabe señalar que se emplearon también algunos otros esquemas de ANN como las redes recurrentes, modelos armax, redes de base radial (RBF) y la red de Kohonen; sin embargo, no se obtuvieron mejoras en los resultados.

Para implementar una ANN hay diversos hiperparámetros como: tipo de red, número de capas de neuronas, número de neuronas en cada capa y la función de activación. Existen algunos trabajos enfocados a buscar la mejor configuración de la ANN. Por ejemplo, Karpenko and Sepehri (2001) realizaron el diagnóstico de fallas en una válvula neumática. Ellos propusieron emplear dos capas ocultas con ocho y cuatro neuronas en cada una y usaron la tangente hiperbólica como función de activación. El número de neuronas fue determinado incrementando el número de neuronas hasta no obtener mejoras en los resultados. Por su parte, Tayarani-Bathaie et al. (2014) realizaron el diagnóstico de fallas en una turbina empleando una ANN con dos capas ocultas. En su trabajo, el número de neuronas para diferentes tipos de fallas fue elegido de manera particular: en un caso usaron cinco neuronas en cada capa, y para otras fallas usaron 20 y 10 neuronas en cada capa.

Aun en el caso de las fallas correctamente aisladas, se requiere esperar algunas muestras para garantizar el umbral alcanzado con el fin de clasificar correctamente la falla presentada. Una manera de mejorar los resultados fue usando dos capas ocultas en la ANN. Si bien, la detección mejoró respecto a lo comentado anteriormente, al incrementar el número de neuronas, se logró una detección superior al 90 % para cada una de las 20 posibles fallas simuladas en el proceso TE. Además de que el error entre la respuesta de la red y la salida deseada se minimizó en el entrenamiento. Desafortunadamente, se incurre en un problema de sobre-modelado (Chen et al., 2014). Lo anterior propicia que no se pueda generalizar la respuesta de la red, aumentando el error al usar datos para validación, con los cuales no se entrenó la ANN.

3.2. Enfoque propuesto de diagnóstico de fallas usando aprendizaje profundo

Antes de alimentar el LSTM se propone realizar un proceso de selección de variables para determinar las variables más representativas y evitar la pérdida de información, ya que el ICA puede verse como una aproximación lineal al proceso TE. Como se ha mencionado, se propone el uso de una red elástica para seleccionar las variables.

3.2.1. Red elástica para la selección de variables

Para determinar cuáles son las variables que reaccionan ante la presencia de fallas, se realiza una selección de variables. La selección de variables puede ser realizada mediante una regresión con penalización usando la norma L_1 en los coeficientes de regresión, resultando así coeficientes contraídos, de donde viene el nombre del algoritmo LASSO. Para emplear el algoritmo se puede utilizar programación cuadrática mediante la regresión de ángulos mínimos (LARS) (Yan y Yao, 2015). En el caso especial de LASSO, se obtienen varios coeficientes igual a cero. De esta manera se obtienen al mismo tiempo, la selección de variables y un modelo de regresión.

La selección de variables se realiza utilizando la optimización de (12) mediante regresión cuadrática

$$\min_B (y - XB)^T (y - XB) + \lambda \|B\|_0, \quad (12)$$

donde y representa la respuesta del sistema, X corresponde a las muestras tomadas del sistema y λ es el multiplicador de Lagrange. Para llevar a cabo la optimización se emplea la norma L_1 (Kuang et al., 2015) obteniendo (13)

$$\min_B (y - XB)^T (y - XB) + \lambda \|B\|_1. \quad (13)$$

La principal limitación del LASSO es tratar variables con alta correlación. Como alternativa, Kuang et al. (2015) propusieron utilizar una red elástica (EN, por sus siglas en inglés), que emplea además la norma L_2 . La EN optimiza a (14)

$$\min_B \left((y - XB)^T (y - XB) + \lambda \frac{1 - \alpha}{2} \|B\|_2^2 + \alpha \|B\|_1 \right), \quad (14)$$

siendo λ y α parámetros a determinar. Se puede notar que si $\alpha = 1$ se tiene el LASSO y además $0 \leq \alpha \leq 1$. En el trabajo de Yan y Yao (2015) se emplean técnicas estadísticas con el LASSO en el proceso TE.

En el presente estudio se seleccionan las variables usando una EN. En la Figura 3 se muestra el resultado obtenido cuando se aplica la EN en la simulación de la falla IDV4. Las variables con las cuales se aproxima mejor el proceso se van eligiendo una a una. De esta manera, la primer variable en este caso x_{28} (componente F al reactor) es la primera en resaltar en la Figura 3.

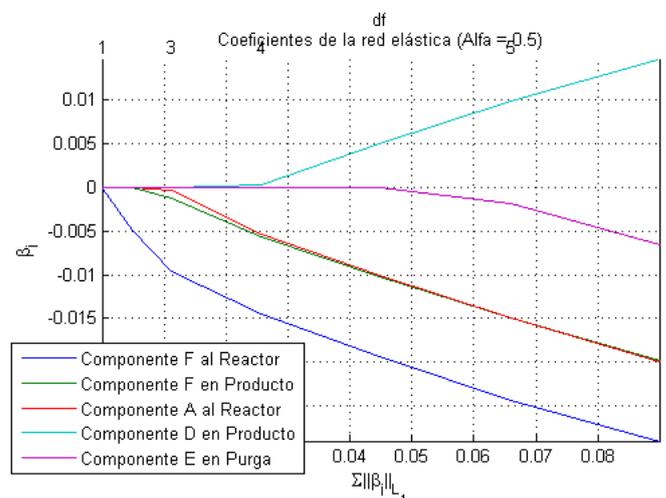


Figura 3: Selección de variables al simular IDV4

Una vez seleccionadas las variables se propone usar una memoria de corto y largo plazo (LSTM) por sus siglas en inglés, la cual forma parte del aprendizaje profundo, siendo esta la mayor aportación. A continuación se muestran las principales contribuciones del aprendizaje profundo en FDI aplicado en el proceso TE, así como la propuesta aquí formada.

3.2.2. LSTM propuesta para diagnóstico de fallas

Las ANN presentan la desventaja de no poder entrenar los pesos sinápticos de las primeras capas debido al desvanecimiento del gradiente cuando se cuenta con más de dos capas ocultas (Wang et al., 2019). Enfrentando este problema, ha surgido un nuevo paradigma conocido como *aprendizaje profundo (DL)*, el cual resulta útil para pre-entrenar una red neuronal profunda mediante el uso de capas sabias. Además, este paradigma ha llegado a considerarse la metodología más efectiva para extraer información (Zhang y Zhao, 2017). Dentro de este paradigma existen varios esquemas que han sido aplicados a diferentes dominios. Una revisión amplia sobre diferentes técnicas de DL en el diagnóstico de fallas en rodamientos de maquinaria es presentada por Hoang y Kang (2019). En su trabajo muestran el uso de autocodificadores apilados (SAE), redes de creencia profundas (DBM) y redes neuronales convolucionales (CNN).

Un trabajo muy completo sobre el uso de DL en el diagnóstico de fallas aplicado al proceso TE fue realizado por Lv et al. (2016). Ellos emplearon autocodificadores apilados. Mientras en (Lv et al., 2017) se utilizaron restricciones dispersas para detectar y remover señales redundantes, para así tener una representación robusta de la información, haciendo uso también de series de tiempo. Por otro lado, en (Lv et al., 2018) se emplean los autocodificadores al proceso TE trabajando en diferentes modos de operación. El uso de redes de creencia profunda para FDI en el proceso TE se muestra en (Zhang y Zhao, 2017). En ese trabajo se seleccionan 10 variables mediante la información mutua (Zheng et al., 2019) y se determinó el número de muestras a procesar para cada tipo de falla. Por su parte, Wu et al. (2018) usaron una CNN con el mismo fin en el proceso TE. Esta estrategia de DL permite una mejor extracción de patrones locales y requiere menor gasto computacional. Además, mediante la desconexión de ciertas neuronas se acelera el aprendizaje y se evita un sobre-ajuste.

En los trabajos mostrados por Lv et al. (2016) se trata de manera especial a las fallas IDV3, IDV9 e IDV15, denominándolas fallas incipientes. Un comentario interesante sobre problemas de aislamiento de fallas se muestra en (Lv et al., 2017), donde se explica la interrelación entre las fallas IDV2, IDV4, IDV9, IDV10, IDV11, IDV13 e IDV14; además en su Tabla 5 se muestra una matriz de confusión entre las fallas IDV4, IDV9 e IDV11.

Las redes neuronales recurrentes tienen la capacidad de reoalimentar información de capas subsiguientes en al menos un ciclo. En el caso particular de las redes con memoria de corto y largo plazo (LSTM) se tiene la capacidad de aprender sobre la dependencia entre términos de corto y largo plazo. Con ello se evita el desvanecimiento del gradiente del error de retropropagación (Fischer y Krauss, 2018) de forma similar a las técnicas de aprendizaje profundo descritas anteriormente.

En (Lv et al., 2017) se propuso emplear series de tiempo para utilizar también la evolución de los datos con el fin de me-

jorar el FDI. En el presente trabajo se propone emplear una red LSTM para contar también con la dependencia entre los datos históricos del proceso medido.

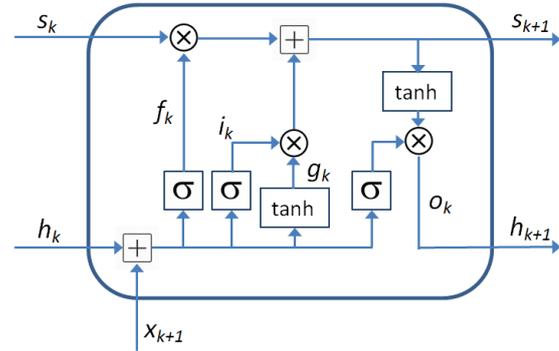


Figura 4: Memoria a corto y largo plazo

Específicamente, las LSTM son redes que cuentan con una capa de entrada cuyo número de neuronas determina el espacio de características. En este tipo de modelos se tienen varias capas ocultas llamadas celdas de memoria y una capa de salida. Esta última capa suele estar completamente conectada y emplea regresión logística en su entrenamiento. La parte más importante de las LSTM son las celdas de memoria, mismas que constan de cuatro compuertas que definen el estado actual s_k de la celda. La primera es la compuerta de olvido f_k encargada de determinar qué información se debe remover de la memoria. Por otro lado, la compuerta de candidatura g_k que define qué información se añadirá a la celda, a su vez la compuerta de entrada i_k determina el control de nivel de actualización del estado en la celda. Finalmente, la compuerta de salida o_k es la encargada de entregar la información de la memoria a la siguiente capa oculta. Los datos que llegan a la celda se denotan como x_k y la información que se envía a las capas ocultas es h_k . Un esquema general de una celda de memoria se muestra en la Figura 4.

El estado siguiente de la celda está dado por (15)

$$s_{k+1} = f_{k+1} \odot s_k + i_{k+1} \odot g_{k+1}, \quad (15)$$

donde \odot representa el producto elemento por elemento (producto de Hadamard).

El valor que se entregará a la siguiente capa oculta se calcula usando (16)

$$h_{k+1} = o_{k+1} \odot \tanh(s_{k+1}), \quad (16)$$

aquí, $\sigma(\cdot)$ representa la función de activación sigmoideal del estado, siendo $\sigma(x) = (1 + e^{-x})^{-1}$.

La actualización de las compuertas se realiza mediante el conjunto (17):

$$\begin{aligned} i_{k+1} &= \sigma(R_i h_k + W_i x_{k+1} + b_i) \\ f_{k+1} &= \sigma(R_f h_k + W_f x_{k+1} + b_f) \\ g_{k+1} &= \tanh(R_g h_k + W_g x_{k+1} + b_g) \\ o_{k+1} &= \sigma(R_o h_k + W_o x_{k+1} + b_o), \end{aligned} \quad (17)$$

donde $\tanh(\cdot)$ es la tangente hiperbólica, W son los pesos sinápticos, R los pesos recurrentes y b los sesgos para cada compuerta. Las compuertas pueden ser vistas como filtros, al inicio la celda determina la información que será olvidada del estado anterior, de manera que si al evaluar la función de activación

$\sigma(\cdot)$ se obtiene un cero, se olvida el dato procesado, no así al obtener un uno. Posteriormente, se determina qué información debe agregarse al estado usando las compuertas de entrada y los valores candidatos. Finalmente, se actualiza el estado y se calcula la salida con la compuerta correspondiente (Fischer y Krauss, 2018).

En el presente trabajo se empleó un abandono (*dropout*) de neuronas con el fin de agilizar el aprendizaje y evitar un sobreentrenamiento. Así, son modificados sólo algunos parámetros en un determinado número de iteraciones, mientras el resto de los pesos sinápticos permanecen constantes. En este sentido, Wu et al. (2018) propusieron una tasa del 50% de abandono en una red neuronal convolucional. En esta propuesta se usó un abandono del 20%, dos capas ocultas con 50 y 30 celdas y una capa final softmax para clasificación con 21 neuronas, una para determinar cada tipo de falla y el estado normal de operación.

Se notó que el FDI es más sensible a hiperparámetros como el tamaño del lote (*batch size*) o la tasa de aprendizaje que con respecto al número de celdas utilizadas, mismas que fueron propuestas al ir incrementando poco a poco la cantidad de las mismas hasta que no se halló una mejora. En este trabajo se usó un minibatch de seis datos con una tasa de aprendizaje de 0.001.

4. Resultados

Para evaluar la detección de fallas, comúnmente se usa el criterio de la tasa de detección de fallas (FDR). Mientras que para evaluar el diagnóstico de fallas se pueden emplear conceptos como la memoria o la precisión, y de manera gráfica, las matrices de confusión. A continuación se muestran los resultados de los dos enfoques planteados.

4.1. Diagnóstico de fallas mediante redes neuronales artificiales multicapa

Para fines de comparación se presentan también los resultados de la ANN jerárquica propuesta en (Xie y Bai, 2015). En la Tabla 1 se muestran los porcentajes FDR obtenidos por la ANN propuesta con los datos preprocesados por ICA y wavelets.

Si bien, los resultados son similares a los reportado en la literatura respecto a la detección de fallas, en una matriz de confusión (Figura 5), se resalta el aislamiento de las fallas usando ANN con dos capas ocultas. En tal representación (matriz de confusión), los resultados ideales concentrarían los círculos de mayor radio en la diagonal. A través de esta matriz, se puede notar que la red tiene dificultad para aislar correctamente las fallas IDV11, IDV12, IDV18 e IDV19 al inferirlas ante la simulación de otras fallas. En contraste, la red reconoce a las fallas IDV1 e IDV6, casi sin proponer la existencia de una falla distinta.

Tabla 1: Tasas de detección de fallas usando ANN

Falla	FDR usando una ANN	FDR (Xie y Bai, 2015)
IDV1	99.41	97.7
IDV2	99.21	97.2
IDV3	93.83	—
IDV4	95.10	99.6
IDV5	98.82	92.8
IDV6	99.41	99.5
IDV7	77.10	99.5
IDV8	97.06	62.3
IDV9	99.41	—
IDV10	99.80	81.8
IDV11	99.02	52.5
IDV12	97.65	55.3
IDV13	99.21	50.4
IDV14	99.41	69.5
IDV15	99.41	—
IDV16	76.12	77.9
IDV17	99.41	71.4
IDV18	99.60	84.3
IDV19	99.21	97.2
IDV20	98.82	73.4

4.2. Diagnóstico de fallas mediante una red LSTM

Las tasas de detección de fallas (FDR) usando DL se muestran en la Tabla 2. A través de la Tabla 2, pueden observarse tasas de detección cercanas al 100% en todos los tipos de fallas. Cabe destacar que el hecho de detectar una falla no implica reconocer qué evento sucedió.

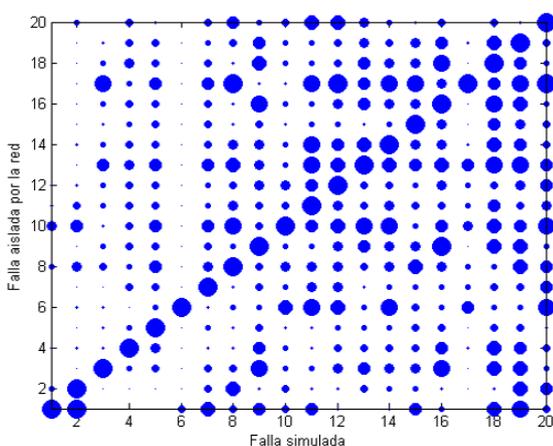


Figura 5: Matriz de confusión usando una ANN

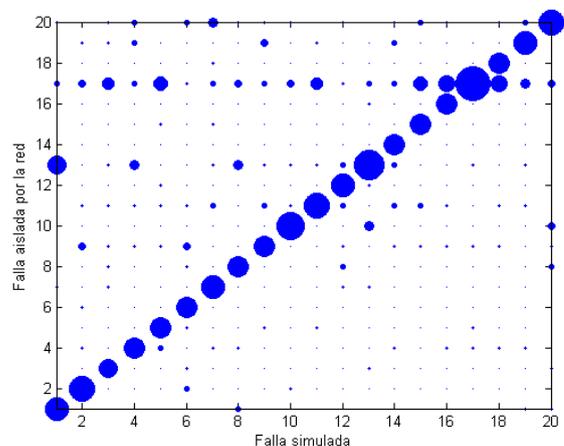


Figura 6: Matriz de confusión usando una LSTM

Tabla 2: Tasas de detección de fallas usando DL en porcentaje

Falla	(Lv et al., 2017)	(Lv et al., 2018)	(Zhang y Zhao, 2017)	(Wu et al., 2018)	LSTM
IDV1	100	100	100	98.6	100
IDV2	99.7	99.7	99	98.5	100
IDV3	95.7	—	95	91.7	96.1
IDV4	100	100	98	97.6	100
IDV5	100	98.8	86	91.5	100
IDV6	100	100	100	97.5	100
IDV7	100	100	100	99.9	100
IDV8	97.8	97.8	78	92.2	100
IDV9	99.2	—	57	58.4	100
IDV10	99.7	99.2	98	96.4	100
IDV11	94.3	89.2	87	98.4	97.8
IDV12	99.7	99.7	85	95.6	100
IDV13	99.7	99.7	88	95.7	100
IDV14	100	95	87	98.7	98.5
IDV15	98.7	—	0	28	98.6
IDV16	99.7	99.5	0	44.2	99.5
IDV17	100	99.7	100	94.5	99.8
IDV18	99.5	99.5	98	93.9	99.6
IDV19	97.8	96.7	93	98.6	99.5
IDV20	99.6	99.3	93	93.3	100

El aislamiento de fallas es reconocer la falla presentada, y para mostrar el desempeño de la LSTM se presenta la matriz de confusión en la Figura 6, donde se aprecian mejores resultados que los mostrados usando una ANN en la Figura 5.

De manera general, para determinar el desempeño de la LSTM en el aislamiento de fallas se presentan también la precisión, memoria y especificidad definidas por (18), (19) y (20) respectivamente (Chen et al., 2014), y mostradas en la Tabla 3.

$$\text{Precisión} = \frac{tp}{tp + fp}, \quad (18)$$

$$\text{Memoria} = \frac{tp}{tp + fn}, \quad (19)$$

$$\text{Especificidad} = \frac{tn}{tn + fp}, \quad (20)$$

donde tp son las mediciones con falla correctamente diagnosticadas como falla (detección de fallas), tn son mediciones de operación normal diagnosticadas como mediciones normales, fp son señales en operación normal diagnosticadas falla (falsas alarmas) y fn es cuando se determina una operación normal cuando existe una falla.

5. Conclusiones

En este trabajo se mostraron los resultados de usar DL para detectar y aislar fallas en el proceso TE. Los resultados son alentadores desde que se obtuvieron tasas de detección de fallas cercanas al 100 % en casi todos los casos. Muchas veces no basta con saber que una falla se ha presentado, sino conocer cuál fue la que se presentó. En este caso, el empleo de una LSTM además de mostrar una FDR alrededor del 100 %, muestra un mejor aislamiento de fallas, comparado contra el uso de una red neuronal clásica como muestran las matrices de confusión. Con respecto a procesar la información más relevante, en el caso de

la red neuronal se empleó el ICA para usar las componentes independientes de mayor magnitud, y para la LSTM se usó una red elástica para seleccionar variables.

Tabla 3: Desempeño del aislamiento de fallas usando LSTM

Falla	Precisión	Memoria	Especificidad
IDV1	96.63	96.38	98.41
IDV2	97.79	93.73	98.97
IDV3	61.50	53.04	78.27
IDV4	66.52	56.33	81.63
IDV5	76.04	59.67	86.84
IDV6	100	99.90	100
IDV7	79.06	65.57	87.81
IDV8	88.74	89.79	88.17
IDV9	72.81	64.14	67.09
IDV10	85.11	82.49	89.82
IDV11	85.92	68.86	90.34
IDV12	85.47	72.77	89.48
IDV13	80.84	59.86	82.87
IDV14	84.87	64.68	84.71
IDV15	73.55	37.22	81.54
IDV16	71.43	55.32	67.38
IDV17	90.12	72.45	86.59
IDV18	82.26	47.82	71.73
IDV19	88.35	57.52	80.83
IDV20	89.78	83.35	89.48

A diferencia de una ANN clásica donde los datos de entrada deben ser escalados para no saturar las funciones de activación, en DL el escalamiento es realizado al entrenarse los pesos sinápticos y sesgos. En este trabajo se empleó una red elástica para entrenar una red con menos parámetros que al haber usado las 41 variables medidas del proceso TE. Cabe señalar que la LSTM también sería capaz de ignorar variables semejantes y el

ruido mediante la compuerta de olvido. Como trabajo futuro, se tiene el interés de detectar y aislar fallas múltiples. Además, se planea comparar el desempeño de la arquitectura propuesta con una arquitectura donde la red elástica sea sustituida por una capa densa regularizando los pesos con las normas L_1 y L_2 . Finalmente, también se desea aplicar técnicas de transferencia de aprendizaje como una herramienta para tratar los tiempos de entrenamiento.

Referencias

- Adewole, A., Tzoneva, R., Behardien, S., 2016. Distribution network fault section identification and fault location using wavelet entropy and neural networks. *Applied Soft Computing* 46, 296–306.
- Alkaya, A., Eker, I., 2011. Variance sensitive adaptive threshold-based PCA method for fault detection with experimental application. *ISA Transactions* 50, 287–302.
DOI: 10.1016/j.isatra.2010.12.004
- Barakat, S., Eteiba, M., Wahba, W., 2014. Fault location in underground cables using anfis nets and discrete wavelet transform. *Journal of Electrical Systems and Information Technology* 1, 198–211.
DOI: 10.1016/j.jesit.2014.12.003
- Bathelt, A., Ricker, N., Jelali, M., 2015. Revision of the Tennessee Eastman process model. *IFAC Papers-Online* 48 (8), 309–314.
DOI: 10.1016/j.ifacol.2015.08.199
- Boldt, F., Rauber, T., Varejao, F., October 2014. Evaluation of the extreme learning machine for automatic fault diagnosis of the Tennessee Eastman chemical process. In: *IEEE (Ed.), Annual Conference of the IEEE Industrial Electronics Society*. Vol. 40. Dallas, Texas, pp. 2551–2557.
DOI: 10.1109/IECON.2014.7048865
- Chen, H., Tino, P., Yao, X., 2014. Cognitive fault diagnosis in Tennessee Eastman process using learning in the model space. *Computers and Chemical Engineering* 67, 33–42.
DOI: 10.1016/j.compchemeng.2014.03.015
- Rodrigues, J., Filho, P., Peixoto Jr., E., Kumar, A., deAlbuquerque, V., 2019. Classification of EEG signals to detect alcoholism using machine learning techniques. *Pattern Recognition Letters* 125, 140–149.
DOI: 10.1016/j.patrec.2019.04.019
- Dixit, A., Majumdar, S., 2013. Comparative analysis of coiflet and daubechies wavelets using global threshold for image denoising. *International Journal of Advances in Engineering & Technology* 6 (5), 2247–2252.
- Downs, J., Vogel, E., 1993. A plant-wide industrial process control problem. *Computers and Chemical Engineering* 17 (3), 245–255.
DOI: 10.1016/0098-1354(93)80018-I
- Fischer, T., Krauss, C., 2018. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research* 270, 654–669.
DOI: 10.1016/j.ejor.2017.11.054
- Gao, X., Hou, J., 2016. An improved SVM integrated GS-PCA fault diagnosis approach of Tennessee Eastman process. *Neurocomputing* 174, 906–911.
DOI: 10.1016/j.neucom.2015.10.018
- Geng, Z., Li, Z., Han, Y., 2018. A new deep belief network based on RBM with glial chains. *Information Sciences* 463, 294–306.
DOI: 10.1016/j.ins.2018.06.043
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press, United States of America, <http://www.deeplearningbook.org>.
- Han, L., Li, C., Guo, S., Su, X., 2015. Feature extraction method of bearing AE signal based on improved Fast-ICA and wavelet packet energy. *Mechanical Systems and Signal Processing* 62-63, 91–99.
DOI: 10.1016/j.ymssp.2015.03.009
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The elements of statistical learning: data mining, inference and prediction*. Springer, New York.
- Hoang, D., Kang, H., 2019. A survey on deep learning based bearing fault diagnosis. *Neurocomputing* 335, 327–335.
DOI: 10.1016/j.neucom.2018.06.078
- Hochreiter, S., Schmidhuber, J., 1997. Long short term memory. *Neural Computation* 9 (8), 1735–1780.
DOI: 10.1162/neco.1997.9.8.1735
- Hyvärinen, A., Oja, E., 2000. Independent component analysis: Algorithms and applications. *Neural Networks* 13, 411–430.
DOI: 10.1016/s0893-6080(00)00026-5
- Jing, C., Gao, X., Zhu, X., Lang, S., July 2014. Fault classification on Tennessee Eastman process: PCA and SVM. In: *IEEE (Ed.), International Conference on Mechatronics and Control*. Jinzhou, China, pp. 2194–2197.
DOI: 10.1109/ICMC.2014.7231958
- Jung, C., Kim, K., Lee, J., Klöckl, B., 2007. Wavelet and neuro-fuzzy based fault location for combined transmission systems. *Energy Systems* 29, 445–454.
DOI: 10.1016/j.ijepes.2006.11.003
- Kandula, V. K., 2011. *Fault detection in process control plants using principal component analysis*. Master's thesis, Louisiana State University, Department of Electrical Engineering.
- Karpenko, M., Sepehri, N., October 2001. A neural network based fault detection and identification scheme for pneumatic process control valves. In: *IEEE (Ed.), International Conference on Systems, Man and Cybernetics*. Tucson, USA, pp. 93–98.
DOI: 10.1109/ICSMC.2001.969794
- Khakipour, M., Safavi, A., Setoodeh, P., 2017. Bearing fault diagnosis with morphological gradient wavelet. *Journal of the Franklin Institute* 354, 2465–2476.
DOI: 10.1016/j.jfranklin.2016.11.013
- Kuang, T., Yang, Z., Yao, Y., 2015. Multivariate fault isolation via variable selection in discriminant analysis. *Journal of Process Control* 35, 30–40.
DOI: 10.1016/j.isatra.2017.06.014
- Kumar, R., Bansal, H., 2019. Hardware in the loop implementation of wavelet based strategy in shuntactive powerfilter to mitigate power quality issues. *Electric Power Systems Research* 169, 92–104.
DOI: 10.1016/j.epsr.2019.01.001
- Lau, C., Ghosh, K., Hussain, M., Hassan, C. C., 2013. Fault diagnosis of Tennessee Eastman process with multi-scale PCA and ANFIS. *Chemom. Intell. Lab. Syst.* 120, 1–14.
DOI: 10.1016/j.chemolab.2012.10.005
- Lee, J., Yoo, C., Lee, I., 2004. Statistical process monitoring with independent component analysis. *Journal of Process Control* 14 (5), 467–485.
DOI: 10.1016/j.jprocont.2003.09.004
- Lei, J., Liu, C., Jiang, D., 2019. Fault diagnosis of wind turbine based on long short-term memory networks. *Renewable Energy* 133, 422–432.
DOI: 10.1016/j.renene.2018.10.031
- Li, W., Monti, A., Ponci, F., 2014. Fault detection and classification in medium voltage DC shipboard power systems with wavelets and artificial neural networks. *IEEE Transactions on Instrumentation and Measurement* 63 (11), 2651–2665.
DOI: 10.1109/TIM.2014.2313035
- Liang, P., Deng, C., Wu, J., Yang, Z., Zhu, J., Zhang, Z., 2019. Compound fault diagnosis of gearboxes via multi-label convolutional neural network and wavelet transform. *Computers in Industry* 113, 103132.
DOI: 10.1016/j.compind.2019.103132
- Lin, J., Zhang, A., 2005. Fault feature separation using wavelet-ICA filter. *NDT&E International* 38, 421–427.
DOI: 10.1016/j.ndteint.2004.11.005
- Linker, R., Gutman, P., Seginer, I., 2002. Observer-based robust failure detection and isolation in greenhouses. *Control Engineering Practice* 10 (5), 519–531.
- Lou, W., Loparo, K., 2004. Bearing fault diagnosis based on wavelet transform and fuzzy inference. *Mechanical Systems and Signal Processing* 18, 1077–1095.
DOI: 10.1016/S0888-3270(03)00077-3
- Lv, F., Wen, C., Bao, Z., Liu, M., 2016. Fault diagnosis based on deep learning. In: *AACC (Ed.), American Control Conference*. Boston, USA, pp. 6851–6856.
DOI: 10.1109/ACC.2016.7526751
- Lv, F., Wen, C., Liu, M., Bao, Z., 2017. Weighted time series fault diagnosis based on a staked sparse autoencoder. *Journal of Chemometrics* 31, 16 pages.
DOI: 10.1002/cem.2912
- Lv, F., Fan, X., Wen, C., Bao, Z., 2018. Stacked sparse auto encoder network based multimode process monitoring. In: *IEEE (Ed.), International Conference on Control Automation & Information Science*. Hangzhou, China, pp. 227–232.
DOI: 10.1109/ICCAIS.2018.8570618
- Maglaveras, N., Stamkopoulos, T., Diamantaras, K., Pappas, C., Strintzis, M., 1998. ECG pattern recognition and classification using non-linear transformations and neural networks: A review. *International Journal of Medical Informatics* 52, 191–208.
DOI: 10.1016/S1386-5056(98)00138-5

- Methnani, S., Lafont, F., Gautier, J., Damak, T., Toumi, A., 2013. Actuator and sensor fault detection, isolation and identification in nonlinear dynamical systems, with applications to a waste water treatment plant. *Journal of Computer Engineering and Informatics* 1 (4), 112–125.
DOI: 10.1080/21642583.2014.888525
- Muñoz-Cobo, J., Mendizábal, R., Miquel, A., Berna, C., Escrivá, A., 2017. Use of the principles of maximum entropy and maximum relative entropy for the determination of uncertain parameter distributions in engineering applications. *Entropy* 19, 486, 37 pages.
DOI: 10.3390/e19090486
- Nguyen, B., Quyen, A., Nguyen, P., Ton, T., July 2017. Wavelet-based neural network for recognition of faults at nhabe power substation of the vietnam power system. In: IEEE (Ed.), International Conference on System Science and Engineering. Ho Chi Minh City, Vietnam, pp. 165–168.
DOI: 10.1109/ICSSE.2017.8030858
- Ojeda-González, A., Mendes-Jr., O., Oliveira-Domingues, M., Menconi, V., 2014. Daubechies wavelet coefficients: a tool to study interplanetary magnetic field fluctuations. *Geofísica Internacional* 53 (2), 101–115.
DOI: 10.1016/S0016-7169(14)71494-1
- Oliveira, J., Pontes, K., Santori, I., Embiruçu, M., 2017. Fault detection and diagnosis in dynamic systems using weightless neural networks. *Expert Systems With Applications* 84, 200–219.
DOI: 10.1016/j.eswa.2017.05.020
- Patan, K., 2008. Artificial neural networks for the modelling and fault diagnosis of technical process. *Lecture Notes in Control and Information Sciences*. Springer, India.
- Rafiee, J., Rafiee, M., Tse, P., 2010. Application of mother wavelet functions for automatic gear and bearing fault diagnosis. *Expert Systems with Applications* 37, 4568–4579.
DOI: 10.1016/j.eswa.2009.12.051
- Ramos-Velasco, L., Ramos-Fernández, J., Islar-Gómez, O., Espejel-Rivera, M., García-Lamont, J., Márquez-Vera, M., 2013. Identificación y control wavelet de un motor de ca. *Revista Iberoamericana de Automática e Informática Industrial* 10, 269–278.
DOI: 10.1016/j.riai.2013.05.002
- Rato, T., Reis, M., 2013. Defining the structure of DPCA models and its impact on process monitoring and prediction activities. *Chemometrics and Intelligent Laboratory Systems* 125, 74–86.
DOI: 10.1016/j.chemolab.2013.03.009
- Rockinger, M., Jondeau, E., 2002. Entropy densities with an application to autoregressive conditional skewness and kurtosis. *Journal of Econometrics* 106, 119–142.
DOI: 10.1016/S0304-4076(01)00092-6
- Salahschoor, K., Kiasi, F., July 2008. On-line process monitoring based on wavelet-ICA methodology. In: IFAC (Ed.), Proceedings of the 17th World Congress. Seul- Korea, pp. 6–11.
DOI: 10.3182/20080706-5-KR-1001.01253
- Salahschoor, K., Khoshro, M., Kordestani, M., 2011. Fault detection and diagnosis of an industrial steam turbine using a distributed configuration of adaptive neuro-fuzzy inference systems. *Simulation Modelling Practice and Theory* 19, 1280–1293.
DOI: 10.1016/j.simpat.2011.01.005
- Sharif, I., Khare, S., 2014. Comparative analysis of Haar and Daubechies wavelet for hyper spectral image classification. In: Commission, I. T. (Ed.), VIII Symposium of The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science. Hyderabad-India, pp. 937–941.
DOI: 10.5194/isprsarchives-XL-8-937-2014
- Smirnov, E., Timoshenko, D., Adrianov, S., 2014. Comparison of regularization methods for imagenet classification with deep convolutional neural networks. *AASRI Procedia* 6, 89–94.
DOI: 10.1016/j.aasri.2014.05.013
- Sobhani-Tehrani, E., Khorasani, K., 2009. Fault diagnosis of nonlinear systems using a hybrid approach. *Fault detection and diagnosis*. Springer, Berlin, Ch. 2, pp. 22–49.
- Tayaran-Bathaie, S., Vanini, Z., Khorasani, K., 2014. Dynamic neural network-based fault diagnosis of gas turbine engines. *Neurocomputing* 125, 153–165.
DOI: 10.1016/j.neucom.2012.06.050
- Žvokelj, M., Zupan, S., Prebil, I., 2016. EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis. *Journal of Sound and Vibration* 26, 394–423.
DOI: 10.1016/j.jsv.2016.01.046
- Wang, X., Qin, Y., Wang, Y., Xiang, S., Chen, H., 2019. ReLTanh: An activation function with vanishing gradient resistance for SAE-based DNNs and its application to rotating machinery fault diagnosis. *Neurocomputing* 363, 88–98.
DOI: 10.1016/j.neucom.2019.07.017
- Wu, F., Tong, F., Yang, Z., 2016. EMGdi signal enhancement based on ICA decomposition and wavelet transform. *Applied Soft Computing* 43, 561–571.
DOI: 10.1016/j.asoc.2016.03.002
- Wu, J., Hsu, C., Wu, G., 2009. Fault gear identification and classification using discrete wavelet transform and adaptive neuro-fuzzy inference. *Expert Systems with Applications* 36, 6244–6255.
DOI: 10.1016/j.eswa.2008.07.023
- Wu, Q., Law, R., Wu, S., 2011. Fault diagnosis of car assembly line based on fuzzy wavelet kernel support vector classifier machine and modified genetic algorithm. *Expert Systems with Applications* 38, 9096–9104.
DOI: 10.1016/j.eswa.2010.12.109
- Wu, H., Zhao., Jinsong, 2018. Deep convolutional neural network model based chemical process fault diagnosis. *Computers and Chemical Engineering* 115, 185–197.
DOI: 10.1016/j.compchemeng.2018.04.009
- Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., Chen, Z., 2019. Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sensing of Environment* 233, 111358.
DOI: 10.1016/j.rse.2019.111358
- Xie, D., Bai, L., December 2015. A hierarchical deep neural network for fault diagnosis on Tennessee-Eastman process. In: IEEE (Ed.), International Conference on Machine Learning and Applications. Vol. 14. Miami, USA, pp. 745–748.
DOI: 10.1109/ICMLA.2015.208
- Yan, R., Gao, R., Chen, X., 2014. Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal Processing* 351, 4555–4569.
DOI: 10.1016/j.sigpro.2013.04.015
- Yan, Z., Yao, Y., 2015. Variable selection method for fault isolation using least absolute shrinkage and selection operator (LASSO). *Chemometrics and Intelligent Laboratory Systems* 146, 136–146.
DOI: 10.1016/j.chemolab.2015.05.019
- Yao, G., Lei, T., Zhong, J., 2019. A review of convolutional-neural-network-based action recognition. *Pattern Recognition Letters* 118, 14–22.
DOI: 10.1016/j.patrec.2018.05.018
- Yin, S., Ding, S., Haghani, A., Hao, H., Zhang, P., 2012. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *Journal of Process Control* 22, 1567–1581.
DOI: 10.1016/j.jprocont.2012.06.009
- Zhang, Q., Yang, L., Chen, Z., Li, P., 2018. A survey on deep learning for big data. *Information Fusion* 42, 146–157.
DOI: 10.1016/j.inffus.2017.10.006
- Zhang, X., Polycarpou, M., Parisini, T., 2002. A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems. *IEEE Transactions on Automatic Control* 47 (4), 576–593.
DOI: 10.1109/9.995036
- Zhang, Y., Zhang, L., Zhang, H., 2012. Fault detection for industrial processes. *Mathematical Problems in Engineering* 2012, 18 pages.
DOI: 10.1155/2012/757828
- Zhang, Z., Zhao, J., 2017. A deep belief network based fault diagnosis model for complex chemical process. *Computers and Chemical Engineering* 107, 395–407.
DOI: 10.1016/j.compchemeng.2017.02.041
- Zhao, H., 2018. Neural component analysis for fault detection. *Chemometrics and Intelligent Laboratory Systems* 176, 11–21.
DOI: 10.1016/j.chemolab.2018.02.001
- Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., Gao, R., 2019. Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing* 115, 213–237.
DOI: 10.1016/j.ymsp.2018.05.050
- Zheng, J., Huang, W., Wang, Z., Liang, J., 2019. Mutual information-based sparse multiblock dissimilarity method for incipient fault detection and diagnosis in plant-wide process. *Journal of Process Control* 83, 63–76.
DOI: 10.1016/j.jprocont.2019.09.004