Anomaly Detection and Diagnosis in Manufacturing Systems: A Comparative Study of Statistical, Machine Learning and Deep Learning Techniques

Kalyani Zope[1], Kuldeep Singh[2], Sri Harsha Nistala[3], Arghya Basak[4], Pradeep Rathore[5] and Venkataramana Runkana[6]

[1,2,3,4,5,6]*TCS Research, Tata Consultancy Services Ltd., Pune – 411013, India*
kalyani.zope@tcs.com
kuldeep.s5@tcs.com
sriharsha.nistala@tcs.com
arghya.basak@tcs.com
rathore.pradeep@tcs.com
venkat.runkana@tcs.com

Abstract

Multivariate sensor data collected from manufacturing and process industries represents actual operational behavior and can be used for predictive maintenance of the plants. Anomaly detection and diagnosis, that forms an integral part of predictive maintenance, in industrial systems is however challenging due to their complex behavior, interactions among sensors, corrective actions of control systems and variability in anomalous behavior. Several statistical techniques for anomaly detection have been in use for a long time and are applied for anomaly detection in continuous as well as batch industrial systems. Further, various machine learning and deep learning techniques for anomaly detection gained significant interest in the recent years. However, anomaly diagnosis that involves localization of the faults did not receive much attention. In this work, we compare the anomaly detection and diagnosis capabilities, in semi-supervised mode, of several statistical, machine learning and deep learning techniques on two systems viz. the interacting quadruple tank (IQT) system and the continuous stirred tank reactor (CSTR) system both of which are representative of the complexity of large industrial systems. The techniques studied include principal component analysis (PCA), Mahalanobis distance (MD), one-class support vector machine (OCSVM), isolation forest, elliptic envelope, dense auto-encoder and long short term memory auto-encoder (LSTM AE). The study revealed that MD and LSTM-AE have the highest anomaly detection capability, followed closely by PCA and OCSVM. The above techniques also exhibited good diagnosis capability. The study indicates that statistical techniques in spite of their simplicity could be as powerful as machine learning and deep learning techniques, and may be considered for anomaly detection and diagnosis in manufacturing systems.

1. Introduction

Manufacturing and process industries such as iron & steel, power, oil & gas, refineries, pharmaceuticals, cement, paper, fine chemicals, etc. comprise of multiple operations and processes that take place in a sequential or parallel manner. These operations involve a wide range of industrial equipment such as furnaces, chemical reactors, steam turbines, gas turbines, bio-reactors, heat exchangers, boilers, condensers, compressors, blowers, fans, pumps, valves, etc. Extensive instrumentation of these industries led to the generation of massive amounts of data from various processes and equipment sensors, mobile and wireless logs, software logs, cameras, microphones and wireless sensor networks at a high frequency (Qin, 2014). This multivariate sensor data represents the true behavior of industrial plants under various operational states such as steady state operation, unsteady state operation, and different operational regimes. For this reason, industrial data can be used to perform predictive maintenance (or condition based maintenance or just-in-time maintenance) of industrial processes and equipment wherein maintenance is scheduled only when the health of the process or equipment falls below a certain threshold, leading to lower cost of inventory and maintenance.

Predictive maintenance involves the following key steps – continuous monitoring of process/equipment using multivariate sensor data, anomaly detection & diagnosis (ADD), estimation of remaining useful life (RUL) of the anomalous unit or component in the plant, and scheduling of maintenance activities based on RULs of anomalous units or components. Anomaly detection refers to identification of atypical patterns in data that do not conform to a notion of normal behavior of a process or equipment. Anomaly diagnosis, also known as fault localization, aims to identify sensors contributing to the detected anomaly and to pinpoint the root cause of the anomaly (Chandola, Banerjee & Kumar, 2009).

Anomalies in industrial data can be categorized into two types viz. 'point anomalies' and 'contextual anomalies'. Point anomalies are those that can be considered anomalous with respect to the rest of the data (e.g. bias in sensor readings) whereas 'contextual anomalies' are those that are anomalous in a specific context but not otherwise (e.g. slow buildup of material in an equipment) (Chandola et al, 2009). Anomaly detection and diagnosis for industrial systems is a challenging task due to the complex behavior of processes and equipment, interactions among sensors,

corrective actions of control systems and variability in anomalous behavior.

A number of data-driven techniques for anomaly detection in industrial data exist in literature. Chandola et al, (2009), surveyed different anomaly detection techniques based on underlying approach adopted by each technique and assessed the effectiveness of these techniques. A similar study by Goldstein and Uchida (2016) evaluated the efficacy of unsupervised anomaly detection algorithms for multivariate data. Qin (2009) applied principal component analysis (PCA) and its variants to identify faults in complex industrial processes and analyzed faults using Hoteling's T^2 and Q statistics. Mujica, Rodellar, Fernandez, and Guem (2010) used PCA for damage assessment in aircraft structures. Chen, Zang, Yuri, Shardt, Ding, Yang, Yang, and Peng (2017) compared the T^2 and Q statistics for fault detection and process monitoring. Soylemezoglu and Jagannathan (2011) proposed Mahalanobis Taguchi System (MTS) based on Mahalanobis Distance (MD) for centrifugal pump failures wherein a fault detection, isolation, and prognostics scheme is presented. They generated fault clusters based on MD values and performed anomaly detection based on thresholds derived from clustering.

Among machine learning techniques, one class support vector machine (OCSVM) is used for detecting anomalies in time-series data by projecting time-series data vectors on to phase spaces (Junshui & Perkins, 2003). Amer, Goldstein, and Abdennadher (2013) proposed OCSVM variants called robust OCSVM and eta-OCSVM that are particularly suited for unsupervised anomaly detection. De Souza, Granzotto, de Almeida, and Oliveira-Lopes (2014) used support vector machine (SVM) for classification and regression, and performed fault detection and diagnosis of a reactor of cyclopentenol production and non-isothermal continuous stirred tank reactor (CSTR) process. Samanazari, Ramezani, Rajabi and Chaibakhsh (2015) used OCSVM for fault detection in a power plant Once-Through Benson Boiler. Liu, Ting, and Zhou (2012) investigated the capability of isolation forest (IF) to handle extremely large data size and high dimensional problems with a large number of irrelevant attributes, and how isolation forest can be used for detection of scattered and clustered anomalies.

Deep learning techniques, especially autoencoders (AE), also demonstrated their effectiveness in unsupervised anomaly detection. Sakurada and Yairi (2014) used a deep autoencoder for feature learning and reconstruction of nonlinear multidimensional data for anomaly detection. Tagawa, Tadokoro, and Yairi (2014) used structured denoising autoencoder for anomaly detection and performed contribution analysis of variables in anomalies. Yan and Yu (2015) used denoising autoencoder for unsupervised representation learning and detected abnormalities in a gas turbine combustor. LSTM autoencoder is also used for anomaly detection in time series data, leveraging its ability to memorize temporal patters in sensory data (Malhotra, Vig, Shroff & Agarwal, 2015).

While there exist plenty of techniques for anomaly detection, techniques for anomaly diagnosis are very limited. Alcala and Qin (2011) analyzed the diagnosability of five PCA-based diagnosis techniques, and proposed and demonstrated the efficacy of 'relative contributions' for better diagnosis on simulated data. Identification of sensors or variables contributing to a detected anomaly is easier than explaining the root causes of anomalies, which usually requires inputs from domain experts. This could be reason for limited attention received by data-driven anomaly diagnosis techniques.

The anomaly detection techniques discussed above have their own distinctive approach to detecting anomalies, and strengths & weaknesses. Statistical techniques such as PCA & MD can handle hundreds of variables but assume linear relationships among variables and are not specifically designed to handle contextual anomalies that occur in temporal industrial data. Machine learning techniques, on the other hand, can model complex nonlinear behaviors but cannot handle too many variables due to their higher computational requirement. Further, deep learning techniques can model relationships that are even more complex and detect contextual anomalies but are computationally intensive for large number of variables.

Each of these techniques is, however, applied to a wide variety of industrial systems and there is no literature, to the authors' knowledge, where the capability of these techniques is compared on an equal footing by using them for anomaly detection on the same system. This motivated us to compare the applicability and capability of data-driven statistical, machine learning and deep learning methods for anomaly detection and diagnosis on the same industrial system. The techniques chosen for the study include PCA, MD, OCSVM, IF, Elliptic Envelope (EE), Dense Autoencoder (Dense-AE) and Long Short Term Memory Autoencoder (LSTM-AE). These techniques are evaluated on two systems viz. the interacting quadruple tank (IQT) and the continuous stirred tank reactor (CSTR), both of which are representative of the complexity of large scale industrial systems.

The rest of the paper is organized as follows. Section 2 describes both the industrial systems and data generation from the same. Section 3 explains the techniques used and the approach adopted for anomaly detection and diagnosis. Results from the experiments are discussed Section 4 and the conclusion is presented in Section 5.

2. INDUSTRIAL SYSTEMS

2.1 Interacting Quadruple Tank (IQT)

The interacting quadruple tank is a well-known benchmark multi-input multi-output (MIMO) system that consists of four interconnected tanks with two pumps and a reservoir

as shown in Figure. 1 (Ozkan, Kara & Arici, 2017). The objective of this process is to maintain the level of the bottom two tanks (Tanks 1 & 2) at desired set points by manipulating the voltage input to both pumps. The inputs to the IQT system are the voltage supply to the pumps, the outputs are the level measurements from the bottom two tanks, and measured variables are the level measurements of all four tanks and input voltage to the system. It is an interacting system where effect of one input is reflected on all other variables. The system has multiple operating points and corresponding responses.

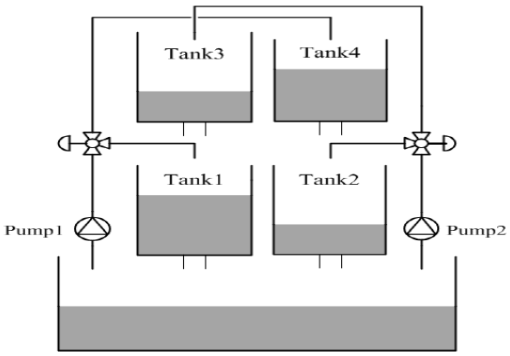

Figure 1. Schematic layout of IQT system

$$A_1 \frac{dh_1}{dt} = -a_1\sqrt{2gh_1} + a_3\sqrt{2gh_3} + \Upsilon_1 k_1 v_1 \quad (1)$$

$$A_2 \frac{dh_2}{dt} = -a_2\sqrt{2gh_2} + a_4\sqrt{2gh_4} + \Upsilon_2 k_2 v_2 \quad (2)$$

$$A_3 \frac{dh_3}{dt} = -a_3\sqrt{2gh_3} + (1-\Upsilon_2) k_2 v_2 \quad (3)$$

$$A_4 \frac{dh_4}{dt} = -a_4\sqrt{2gh_4} + (1-\Upsilon_1) k_1 v_1 \quad (4)$$

where,
A_i: cross-sectional area of i^{th} tank
h_i: water level of i^{th} tank
a_i: cross-sectional area of the outlet pipes of i^{th} tank
Υ_i: value ratio of i^{th} valve position
k_i: pump constant of i^{th} pump

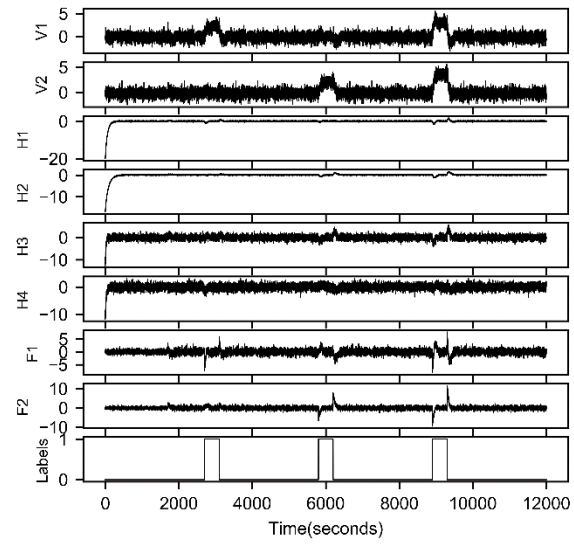

Figure 2. Trends of variables in IQT system showing normal & faulty operation

For this study, the model equations of the IQT system (Eqs. 1 to 4) are simulated using Xcos (Scilab), an open source software environment used for modelling & simulation, to generate operational data. Frictional losses in the IQT system are considered to be negligible. Normal operation (consisting of start-up, set point up and down, shut down) as well as faulty operation of the system are simulated.

A total of 11 different process operations were simulated and corresponding data consisting of 8 variables is recorded at sampling frequency of 1/sec for a duration of 3 hours. The response time of the IQT system is ~1.5 secs and the sampling frequency of 1/sec is sufficient to capture the system dynamics in both normal and faulty operation. Out of the 11 scenarios, three correspond to faulty operation while eight correspond to normal operation. Normal operation includes operation at various normal modes/regimes that occur due to changing set points. The faults introduced in the system include clogging of valves and leakage in one of the tanks. Figure. 2 shows the trends of scaled data from the IQT variables for a period of 3 hours. V_1 and V_2 are the input voltages of Pump1 and Pump2 respectively. H_1, H_2, H_3 and H_4 are the levels of four tanks, and F_1 and F_2 are the water flow rates. 'Labels' indicates the type of operation with '0' and '1' referring to normal and faulty operation respectively.

2.2 Continuous Stirred Tank Reactor (CSTR)

The continuous stirred reactor system (Martin, 1995), forms an integral part of several manufacturing and chemical process industries such as oil and gas, fine chemicals, etc. The system studied here comprises of a CSTR in which a generalized catalytic first order exothermic reaction A → B takes place. The process is non-isothermal and the temperature in the reactor is maintained by indirect heat exchange with cool water flowing through a cooling coil. The system has one input feed stream, which is a mixture of solute (F_A) and solvent (F_S), one output stream emanating from the reactor and one water stream (F_C) as shown in Figure. 3 (Yoon & MacGregor, 2001).

This system is also a MIMO system with highly nonlinear system dynamics. From a control perspective, the objective of the process is to maintain the composition and temperature of the reactor outlet stream. The outlet stream composition and temperature can be maintained by modulating the manipulated variables viz. F_C and F_A respectively. A conventional PID control strategy is used to control the yield and temperature of the system. Both the control loops interact with one another, making the system more difficult to control. The system also has several measured disturbances such as solute concentration (C_{AA}), solvent composition (C_{AS}), solute flow rate (F_S), temperature of the reactor inlet stream (T_0) and temperature of inlet water flow rate (T_C) as well as unmeasured disturbances such as poisoning of the catalyst and fouling of the cooling coil.

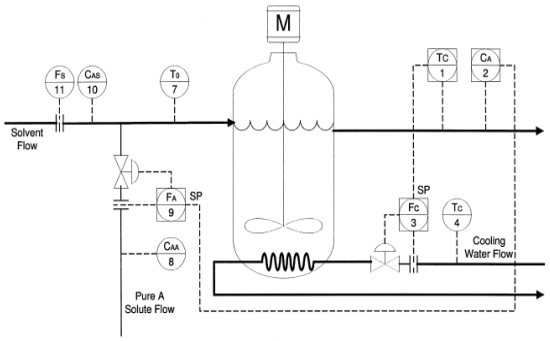

Figure 3. Schematic layout of CSTR with feedback control system

For this study, the model equations of CSTR system (Eqs. 5 to 6) are simulated using Xcos for 14 different operating regimes and corresponding data consisting of 7 variables is generated with a sampling frequency of 1/sec for a total duration of 2.5 hours. The response time of the IQT system is ~ 1 sec and the sampling frequency of 1/sec is sufficient to capture the system dynamics in both normal and faulty operation. Data is generated for normal operation as well as faulty operation. While normal operation consists of start-up, shut down, and set point change by the operator, faulty operation consists of initiation and propagation of faults such as bias in outlet temperature sensor measurement (F#1), bias in inlet temperature sensor measurement (F#2), bias in inlet reactant concentration (F#3), catalyst poisoning (F#4) and cooling coil fouling (F#5).

$$\frac{dC_A}{dt} = \frac{F}{V}C_{A0} - \frac{F}{V}C_A - k_0 e^{-\frac{E}{RT}}C_A \quad (5)$$

$$V\rho c_p \frac{dT}{dt} = \rho c_p F(T_0 - T) - \frac{aF_c^{b+1}}{F_c + \frac{aF_c^b}{2\rho_c c_{pc}}}(T - T_{c,in}) + (-\Delta H_{rxn})Vk_0 e^{-\frac{E}{RT}}C_A \quad (6)$$

where,
C_A, C_{A0}: outlet and inlet concentrations of the species 'A'
F, F_c: molar flow rates of inlet stream and coolant
V: volume of the reaction mixture in the tank
k_0: pre-exponential rate kinetics constant
E: activation energy of the reaction
R: universal gas constant
T, $T_{c,in}$: reaction and coolant temperatures
ρ, ρ_c: densities of the mixture and coolant
c_p, c_{pc}: specific heat capacities of the reaction mixture and coolant
ΔH_{rxn}: heat of reaction

Figure. 4 shows the trends of scaled data from the CSTR variables for a period of 2.5 hrs. C_{in} and C_{out} are measured inlet and outlet concentrations, T_{in} and T_{out} are inlet and outlet temperatures, T_c is coolant (water) temperature, F_s is flow rate of solute A and F_c is flow rate of water.

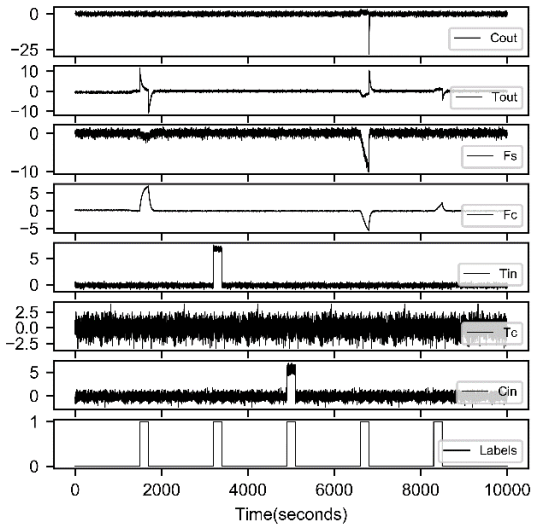

Figure 4. Trends of variables in CSTR system showing normal & faulty operation

3. ANOMALY DETECTION & DIAGNOSIS

As mentioned in Section 1, seven techniques viz. PCA, MD, OCSVM, IF, EE, Dense-AE and LSTM-AE are used for anomaly detection in this study. Each of the techniques is used in a semi-supervised mode, i.e. the technique is used to model only the normal behavior of the system, where it has worked smoothly without any faults, to obtain an 'anomaly detection model'. This model is then used to test if any other operational data falls within or outside the threshold of normal behavior by means of an anomaly score. Anomaly diagnosis, i.e. identification of sensors contributing to the anomaly is carried out only for anomalous operational data. Diagnosis is performed using three of the techniques viz. PCA, MD & LSTM-AE. For each technique, the score used for detecting anomalies and the method used for diagnosing the anomalies are shown in Table 1.

Table 1. Anomaly Scores and Diagnosis Methods

Technique	Score used for Anomaly Detection	Anomaly Diagnosis Method
PCA	Hoteling's T^2	Complete decomposition contributions (CDC) to T^2
MD	MD	One-variable substitution method
OCSVM	Distance to separating hyperplane	-
IF	Average path length from root node	-
EE	Negative Mahalanobis Distance	-
Dense-AE	Reconstruction Error (RE)	-
LSTM-AE	RE	RE of individual sensor

For PCA technique, the Hoteling's T^2 statistic is used as the anomaly score and the complete decomposition contributions (CDC) to T^2 are used for anomaly diagnosis (Alcala & Qin, 2011). For MD technique, the Mahalanobis distance itself is used as the anomaly score. A novel method of substituting each variable from the anomalous instance

into the nearest normal instance, recalculating the MD and identifying the variables with the highest substituted MD as 'contributors to the anomaly' is used for anomaly diagnosis.

The OCSVM technique maps the original data into a high dimensional space and learns a decision function to represent the normal data. A hyperplane in the mapped space separates the majority of the data from the origin and points that lie on the other side of the hyperplane are identified to be anomalous. OCSVM therefore uses the distance of each point to the hyperplane as the anomaly score.

IF is an isolation-based method and works differently from methods that employ distance or density based metrics. It isolates instances by randomly selecting a feature and split value between the range of selected feature. Path length is the number of splits required to isolate a sample from root node to leaf node. This path length is shorter for anomalous data compared to that for normal data. Hence, when a forest of random trees collectively produces shorter path lengths for particular samples, they are classified as anomalies. Average of path lengths of an input vector from root node in all trees in the forest is used as the anomaly score (Liu et al, 2012).

The EE technique assumes that the data follows a known distribution (e.g. Gaussian distribution). It finds the possible covariance between feature dimensions in multivariate data and fits an ellipse to the central data points. The data points far enough from the fitted ellipse are classified as anomalies. Negative Mahalanobis distance of input vector from ellipse boundary is used as the anomaly score (Rousseeuw and Driessen, 1999).

Autoencoder is an artificial neural network composed of two feed-forward neural networks known as encoder and decoder where the input layer of encoder and output layer of decoder have the same dimension. While the encoder learns the encodings of input data and transforms it into a latent space, the decoder brings back data from the latent space to the original data space. This procedure is known as reconstruction of input data. Long short-term memory networks are special type of recurrent neural networks (RNNs) widely used for time series data. A traditional neural network considers all instances of input data to be independent of each other and cannot make use of sequential information. LSTMs have an edge over conventional feed-forward neural networks in remembering long term dependencies in data by making use of previous hidden state information in addition to the current input provided through input layer. In LSTM, the information flows through a mechanism known as cell states. The flow of information in memory cell is regulated by input, output and forget gates.

In the Dense-AE technique, regular neurons are used in the autoencoder while in the LSTM-AE technique, LSTM cells are used in the autoencoder. In both the autoencoders, reconstruction error (RE) of input data, i.e., squared difference between reconstructed data and original data is used as the anomaly score. For anomaly diagnosis, the reconstruction error of individual sensors is compared against the corresponding thresholds. Sensors whose RE exceeds their corresponding thresholds are identified as anomalous sensors (Zope, Nistala & Runkana, 2018). The anomaly detection and diagnosis approach in this work consists of two phases viz. training phase and testing phase.

3.1 Training Phase

The sequence of steps followed in the training phase are shown in Figure. 5.

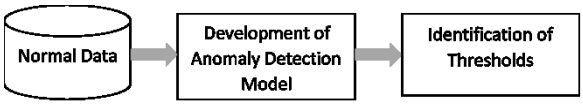

Figure 5. Sequence of steps in the training phase

Selection of Data

For both the systems, normal data where the system has worked smoothly without any anomalies is considered as training data to build the anomaly detection model. 20% of training data is used as validation data for determining of thresholds of the anomaly scores. The data apart from training data is designated as test data. The training data is first subjected to data pre-processing, where start-up and shutdown periods are removed, and sensor measurements are z-normalized using the mean and standard deviation of the training data. For the IQT system, the number of training instances is 4900, and test data consists of 5400 normal instances and 1200 faulty instances. For the CSTR system, the number of training instances is 4500, and test data consists of 4500 normal instances and 1000 faulty instances.

Development of Anomaly Detection Model

As mentioned earlier, the normal operational data from each system is used for building the anomaly detection models. These models are of the form

$$Y_i = f_k(X_i) \tag{7}$$

where

Y_i is the anomaly score

X_i is the input data vector

f_k is the functional form linking to X_i to Y_i for the k^{th} anomaly detection technique

Hyper-parameter tuning is performed for the techniques to obtain optimal values of hyper-parameters and the best possible anomaly detection models. For PCA, the number of principal components explaining 95% of variance is considered. For OCSVM, RBF kernel was used. Parameters, gamma & nu are tuned in the ranges 1^{-9} to 1^{3} and 0.01 to 0.2 (incremented by a factor of 10) respectively. In IF, the number of base estimators is tuned in the range 80 to 200 in increments by 10; the proportion of outliers (contamination) in the dataset was optimized in the range

0.1 to 0.20. For EE, the proportion of outliers (contamination) is optimized in the range 0.05 to 0.30.

In the Dense-AE, three dense layers are stacked in the encoder and the decoder. Similarly, in the LSTM-AE, three LSTM layers are stacked in the encoder and the decoder. Learning rate, batch size and window size (for LSTM-AE only) are the hyper-parameters considered while tuning AE models. Learning rate represents the rate at which weights in the network are updated. Learning rates of 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5} are used. Considering the dynamics of the systems and data size, batch sizes are set at 32, 100 and 250, and the window size (number of time steps in each window) is set at 4, 10, 20 and 30 for LSTM autoencoder. Batch size is number of instances used at once for weight updation. Each AE model is trained for a maximum of 300 epochs. The mean squared error of reconstruction is computed after each epoch and training is terminated if the mean squared error did not improve for certain number of epochs (early stopping) to prevent overfitting on training data. The optimal values of hyper-parameters obtained for various anomaly detection models are shown in Table. 2.

Table 2.Optimal hyper-parameters for anomaly detection models

Technique	Optimal Hyper-parameters	
	IQT System	CSTR System
PCA	# of PCs: 7	# of PCs: 6
MD	None	None
OCSVM	Gamma: 0.1 Nu: 0.11	Gamma: 0.1 Nu: 0.01
IF	# of estimators: 90 Contamination: 0.15	# of estimators: 80 Contamination: 0.15
EE	Contamination: 0.1	Contamination: 0.05
Dense-AE	Batch size: 32 # of neurons: 4	Batch size: 32 # of neurons: 2
LSTM-AE	Window size: 4 Batch size: 32 # of LSTM cells: 250, 128	Window size: 30 Batch size: 32 # of LSTM cells: 250, 250

Identification of Anomaly Score Thresholds

For the PCA technique, the threshold for T^2 statistic is obtained from the Chi^2 test. For MD, OCSVM, IF, EE, Dense-AE and LSTM-AE techniques, the thresholds for anomaly scores are estimated as $\mu + k\sigma$ where μ and σ are the mean and standard deviation of anomaly score for the validation data, and k is a real number chosen such that anomaly scores for 99% of validation data is within $\mu + k\sigma$.

3.2 Testing Phase

The sequence of steps followed in the testing phase are shown in Figure 6.

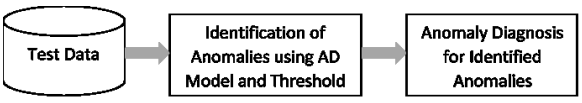

Figure. 6 Sequence of steps in the testing phase

Anomaly Detection

The anomaly detection models built during the training phase are used for detecting anomalies in the test data that consists of normal as well as anomalous data instances. For each technique, the anomaly score obtained for each instance is compared against the corresponding anomaly score threshold. Instances for which the anomaly scores exceeds the corresponding threshold are categorized as 'anomalous'.

Anomaly Diagnosis

Anomaly diagnosis is performed for each anomalous instance to identify the subset of sensors contributing to the anomaly. The methods shown in Table 1 are used to conduct anomaly diagnosis using PCA, MD and LSTM-AE techniques. Results from anomaly diagnosis are compared against the known 'ground truth' to evaluate the efficacy of various diagnosis methods.

4. Results & Discussion

For both the IQT and CSTR systems, receiver operating characteristic (ROC) curves for each anomaly detection technique are obtained by varying the anomaly score thresholds and computing the true positive rate (TPR) and false positive rate (FPR) on the test data. Each point on the ROC curve represents a classifier (i.e. anomaly detection model with a specific anomaly score threshold) with the corresponding TPR & FPR values; the classifier corresponding to TPR ~ 1 and FPR ~ 0 (top left corner) is the perfect classifier. The area under the ROC curve (AUC) is a measure of the efficacy of classification performance (in this case, anomaly detection performance). The AUC is higher for the ROC curves approaching the perfect classifier.

The ROC curves for the IQT and CSTR systems are shown in Figure. 7 & 8 respectively. The AUC values (expressed as a percentage) for each anomaly detection technique and both the systems are shown in Table 3. From Figure. 7 and Table 3, it can be observed that for the IQT system, LSTM-AE demonstrated the highest anomaly detection capability followed by PCA, MD and OCSVM with almost similar AUC values. From Figure. 8 and Table 3, it can be observed that for the CSTR system, MD & EE demonstrated the highest anomaly detection capability, followed closely by LSTM-AE and OCSVM.

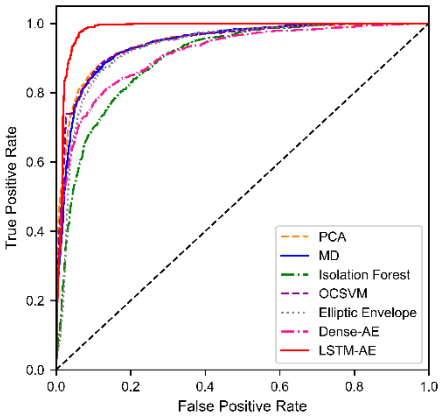

Figure. 7 ROC curves for anomaly detection techniques for IQT system

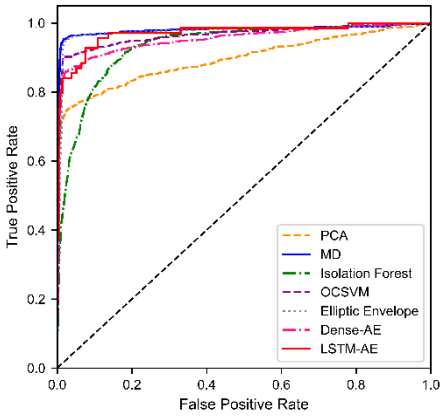

Figure. 8 ROC curves for anomaly detection techniques for CSTR system

Table 3. AUC Scores

System	PCA	MD	OCSVM	IF	EE	Dense-AE	LSTM-AE
IQT	94.8	94.2	94.6	89.8	93.1	91.3	98.3
CSTR	89.5	98.0	96.4	93.4	97.5	95.3	97.1

Color Code: Highest AUC, 2nd Highest AUC, 3rd Highest AUC, 4th Highest AUC

Various anomaly detection metrics viz. accuracy, precision, recall, F1-score, false positive rate and missed detection rate (MDR) for the anomaly detection models built during the training phase are shown in Tables 4 and 5 for IQT and CSTR systems respectively. MDR and FPR are key metrics for industrial anomaly detection models. MDR refers to the number of actual faults that are not detected and should be as low as possible, particularly for industrial systems that need to be tightly controlled or prone to runaway faults with serious consequences. Here, it may be noted that MDR = 1 – TPR (or 1 – Recall). FPR refers to faults that are detected but do not exist (i.e. false alarms). Higher FPR mandates operators to take corrective action when it is not needed that may drive the system in an undesirable direction. Therefore, FPR should also be as low as possible (Tien, Lim & Jun, 2004). Other metrics such as accuracy, precision, recall and F1-score should be as high as possible.

Table 4. Anomaly detection metrics for IQT system

Technique	Accuracy	P	R	F1-score	FPR	MDR
PCA	0.92	0.77	0.78	0.78	0.05	0.22
MD	0.90	0.59	0.83	0.69	0.09	0.17
OCSVM	0.93	0.74	0.83	0.78	0.06	0.17
IF	0.83	0.78	0.52	0.63	0.05	0.48
EE	0.84	0.91	0.53	0.67	0.02	0.47
Dense-AE	0.89	0.50	0.85	0.63	0.10	0.15
LSTM-AE	0.86	0.26	0.93	0.41	0.14	0.07

P: Precision, R: Recall (or TPR), MDR: Missed Detection Rate (= 1 – TPR)
Color Code: Best value of metric, 2nd best value of metric, 3rd best value of metric

From Table 4, it can be observed that for the IQT system, LSTM-AE has the lowest MDR, followed by Dense-AE, MD and OCSVM, while IF and EE have highest MDR. In terms of FPR, EE has the lowest value, followed by IF, PCA and OCSVM. Figure. 9 depicts the MDR values for the 7 techniques plotted against the corresponding FPR values. From the figure, it can be observed that MD, OCSVM and Dense-AE techniques fall in the quadrant with acceptable MDR (< 0.2) & FDR (< 0.1) thresholds. It may noted that the three techniques have high values of accuracy and F1-score and reasonable values of precision.

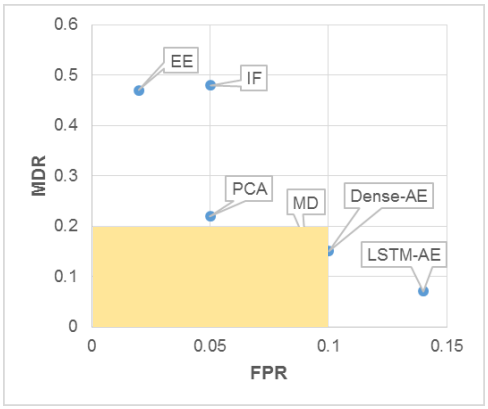

Figure. 9 MDR vs FPR for IQT system (yellow box indicates quadrant with acceptable thresholds)

For the CSTR system, it can be observed from Table 5 that Dense-AE has the lowest MDR followed by LSTM-AE, OCSVM and MD while IF and EE once again have the highest MDR. In terms of FPR, EE and MD have the lowest values followed by OCSVM. Fig. 10 depicts the MDR values for the 7 techniques plotted against the corresponding FPR values for this system. It can be observed from the figure that MD, PCA, OCSVM, Dense-AE and LSTM-AE fall in the quadrant with acceptable MDR (< 0.2) & FDR (< 0.06) thresholds. Further, these techniques have high values of accuracy, F1-score and precision.

Table 5. Anomaly detection metrics for CSTR system

Technique	Accuracy	P	R	F1-score	FPR	MDR
PCA	0.92	0.76	0.81	0.78	0.05	0.19
MD	0.96	0.96	0.85	0.90	0.01	0.15
OCSVM	0.96	0.85	0.93	0.89	0.03	0.07
IF	0.86	0.81	0.59	0.68	0.05	0.41
EE	0.85	0.97	0.55	0.70	0.01	0.45
Dense-AE	0.95	0.77	0.95	0.85	0.05	0.05
LSTM-AE	0.95	0.80	0.93	0.86	0.05	0.07

P: Precision, R: Recall (or TPR), MDR: Missed Detection Rate (= 1 – TPR)
Color Code: Best value of metric, 2nd best value of metric, 3rd best value of metric

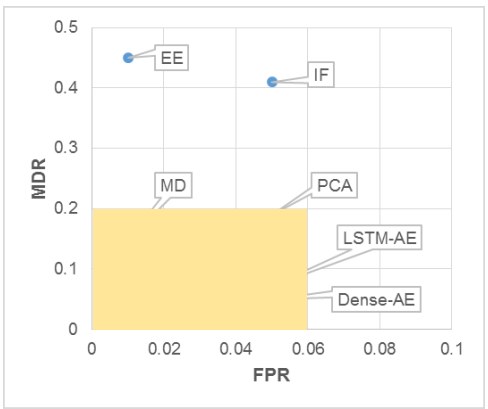

Figure. 10 MDR vs FPR for CSTR system (yellow box indicates quadrant with acceptable thresholds)

The trend of anomaly scores obtained from the anomaly detection models from each of the techniques for the IQT and CSTR systems are shown in Fig. 11 & 12 respectively. The 'Labels' plot in Fig. 9 & 10 indicates the ground truth where the label value of '0' indicates normal data and the label value of '1' indicates anomalous data.

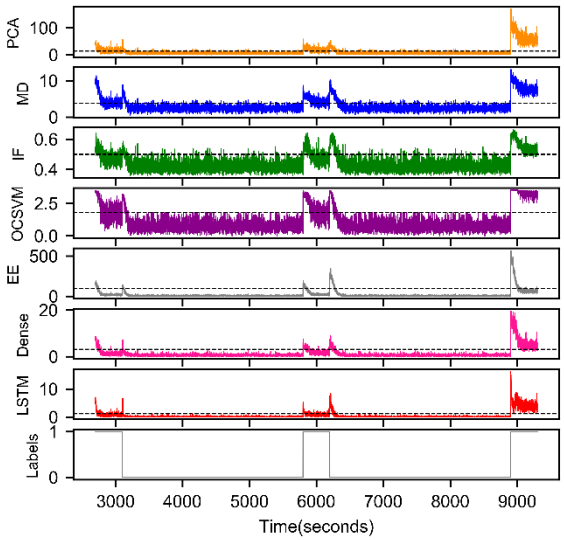

Figure. 11 Trend of anomaly scores from different anomaly detection techniques for IQT system (dotted lines indicates anomaly score thresholds)

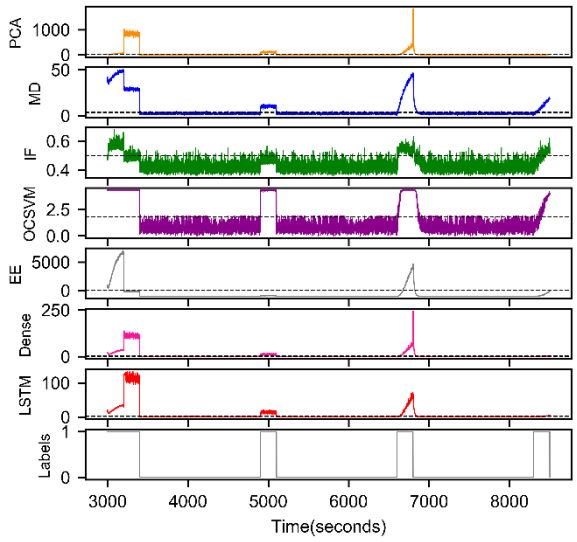

Figure. 12 Trend of anomaly scores from different anomaly detection techniques for CSTR system (dotted lines indicates anomaly score thresholds)

From Fig. 11, it can be observed that the anomaly scores from all the techniques crossed their respective thresholds at the anomalous events around time periods of 3000, 6000 and 9000. This indicates that all the techniques are able to identify the anomalies introduced in the IQT system. It should, however, be noted that the separation between scores for normal & faulty instances and number of false positives (score crossing the threshold even for normal instances) are different for different techniques. On similar lines, it can be observed from Fig. 12 that the anomaly scores from almost all the techniques crossed their respective thresholds at the anomalous events around time periods of 3200, 5000, 6700 and 8200. This indicates that the techniques are able to identify the anomalies introduced into the CSTR system. However, in this case as well, the separation between scores for normal & faulty instances and number of false positives is different for different techniques.

The performance of anomaly diagnosis methods associated with PCA, MD and LSTM-AE techniques for IQT and CSTR systems is summarized in Tables 6 & 7 respectively. From Table 6, it can be observed that for each fault in the IQT system, the key variables (top 2 or top 3 variables) identified to be responsible from the three techniques are in line with the ground truth. It should be noted that apart from the variables actually responsible for the anomaly, the diagnosis techniques identified few additional variables. This may be attributed to the interactions among the variables due to which the effect of original faulty variables gets reflected in various other variables. Similarly, from Table 7, it can be observed that the three diagnosis methods correctly identified the key variables responsible for each of the faults introduced in the CSTR system. This indicates that the three techniques have strong diagnosis capability w.r.t faults in industrial systems.

Table 6. Diagnosis of faults in IQT system

Fault	Period of fault	Variables indicative of fault (Ground Truth)	Anomaly Diagnosis	
			Technique	Variables identified
Clogging of Valve 1	2700 to 3100	V1	PCA	V1, (F1)*
			MD	V1, (F1)
			LSTM-RE	V1, (F1)
Clogging of Valve 2	5800 to 6200	V2	PCA	V2, (F2)
			MD	V2, (F2)
			LSTM-RE	V2, (F2)
Leakage in Tank 1	8900 to 9300	V1, V2	PCA	V1, V2, (F1, F2, H4)
			MD	V1, V2, (F1, F2, H4)
			LSTM-RE	V1, V2, (F2)

*Variables shown within ( ) are additional variables in which fault signature is reflected

Table 7. Diagnosis of faults in CSTR system

Fault	Period of fault	Variables indicative of fault (Ground Truth)	Anomaly Diagnosis	
			Technique	Variables identified
Bias in Outlet Temperature	1500 to 1700	Fc	PCA	Fc, (Tout)
			MD	Fc, (Tout)
			LSTM-AE	Fc, (Tout)
Bias in Inlet Temperature	3200 to 3400	Tin	PCA	Tin, (Cout)
			MD	Tin, (Cout)
			LSTM-AE	Tin
Bias in Inlet Concentration	4900 to 5100	Cin	PCA	Cin, (Cout)
			MD	Cin, (Cout)
			LSTM-AE	Cin
Catalyst Poisoning	6600 to 6800	Fc, Fs	PCA	Fc, Fs, (Tout)
			MD	Fc, Fs, (Tout)
			LSTM-AE	Fc, Fs, (Tout)
Cooling Coil Fouling	8300 to 8500	Fc, Tout	PCA	Fc, Tout
			MD	Fc, Tout
			LSTM-AE	Fc, Tout

*Variables shown within ( ) are additional variables in which fault signature is reflected

The comparative study of statistical, machine learning and deep learning based anomaly detection techniques revealed that there is no single best technique for all industrial systems. This is understood given the fact that industrial systems exhibit a wide range of operational behavior and experience simple as well as complex anomalies. Further, the efficacy of anomaly detection depends, other than the chosen technique, on the number and location of sensors used, the quality of training data, the severity of anomalies and the extent of corrective control action on the system. However, by applying all the anomaly detection techniques on the same systems, it is identified that simpler statistical techniques such as MD possess comparable anomaly detection capability as that of machine learning techniques (e.g. OCSVM) and deep learning techniques (e.g. LSTM-AE). This is the key outcome of this study. Further, anomaly diagnosis methods based on MD, PCA and LSTM-AE provided similar results. It should be noted here that none of the techniques, except LSTM, explicitly learnt the temporal patterns (i.e. the techniques consider the data points to be unrelated in time) and yet are able to successfully detect contextual anomalies in the two systems considered in this study. The ability of these techniques to detect and diagnose contextual anomalies needs to be tested further using other complex industrial systems, possibly involving multiple operating regimes.

Statistical anomaly detection models (e.g. MD or PCA-based models) can be developed with little computational effort. They can be built using limited data as well as large amounts of data, and are highly scalable (i.e. they can be easily extended to hundreds of variables). However, such techniques typically assume linear relationships among variables (though there are nonlinear variants of PCA) and may not perform well when applied to industrial systems involving highly nonlinear dynamics. On the other hand, ML/DL based anomaly detection models (tree, kernel or neural network-based models) can be used to learn nonlinear behavior of industrial systems. However, they require considerable computational effort and resources (e.g. GPUs for training deep learning models). They also necessitate large amounts of data to prevent overfitting and for learning hyper-parameters. Majority of the time, more data leads to improved generalizability of ML/DL models. Further, ML/DL models do not scale easily as large number of variables lead to large computational time.

In this context, it may be advantageous to combine the strengths of statistical techniques with those of ML/DL techniques for anomaly detection and diagnosis in manufacturing and process industries that involve large number of variables and exhibit nonlinear behavior. Statistical techniques can be used either to corroborate the results from other complex techniques or in conjunction with ML/DL techniques in an 'ensemble' anomaly detection and diagnosis framework.

5. CONCLUSION

In this work, we compared the capabilities of various statistical, machine learning and deep learning techniques for anomaly detection and diagnosis on an interacting quadruple tank system and a continuous stirred tank reactor system. We studied the anomaly detection performance of PCA, MD, OCSVM, IF, EE, Dense-AE and LSTM-AE in semi-supervised mode. Empirical studies on two industrial datasets demonstrated that MD and LSTM-AE have the highest anomaly detection capability, followed by PCA and OCSVM. For detected anomalies, variables responsible for each of the anomalies are diagnosed using PCA, MD and LSTM-AE techniques. All three techniques provide very similar diagnosis results on both datasets. This study helped conclude that statistical anomaly detection and diagnosis techniques deliver results comparable to more complex ML/DL techniques, and may be therefore considered alongside ML/DL techniques for anomaly detection and diagnosis in manufacturing and process industries.

REFERENCES

Alcala, C. F., & Qin, S. J. (2010). "Reconstruction-based contribution for process monitoring with kernel principal component analysis". *Proceedings of the 2010 American Control Conference*, 29 July 2010, Baltimore, MD, USA, DOI: 10.1109/ACC.2010.5531315

Alcala, C. F., & Qin, S. J. (2011). "Analysis and generalization of fault diagnosis methods for process monitoring". *Journal of Process Control* - J PROCESS CONTROL. 21. 322-330. 10.1016/j.jprocont.2010.10.005.

Amer, M., Goldstein, M. & Abdennadher, S. (2013). "Enhancing one-class support vector machines for unsupervised anomaly detection". *Proceedings of the ACM SIGKDD Workshop on Outlier Detection and Description,* August 11-11, New York, NY, USA, DOI: 10.1145/2500853.2500857

Chandola, V., Banerjee, A. & Kumar, V. (2009). "Anomaly Detection: A Survey". *ACM Computing Surveys, 1-72*

Chen, Z., Zhang, K., Yuri A.W., Shardt, Ding, S. X., Yang, X., Yang C. & Peng, T. (2017). "Comparison of Two Basic Statistics for Fault Detection and Process Monitoring". *International Federation of Automatic Control*, Volume 50, Issue 1, July 2017, Pages 14776-14781

Choi, S. W., Lee, C., Lee, J. M., Park, J. H., & Lee, I. B. (2005). "Fault detection and identification of nonlinear processes based on kernel PCA". *Chemometrics and intelligent laboratory systems,* 75(1), 55-67, DOI: 10.1016/j.chemolab.2004.05.001.

De Souza, D. L., Granzotto, M. H., De Almeida, G. M. & Oliveira-Lopes, L. C. (2014). "Fault Detection and Diagnosis Using Support Vector Machines - A SVC and SVR Comparison". *Journal of Safety Engineering*, 2014, 3(1): 18-29 DOI: 10.5923/j.safety.20140301.03

Goldstein, M. & Uchida, S. (2016). "A comparative evaluation of unsupervised anomaly detection

algorithms for multivariate data*"*. April 19, 2016, DOI:10.1371/journal.pone.0152173.

Hoyle, B., Rau, M. M., Paech, K., Bonnett, C., Seitz, S. & Weller, J. (2016). "Anomaly detection for machine learning redshifts applied to SDSS galaxies". arXiv:1503.08214 [astro-ph.CO], Jun 15 2016, DOI:10.1093/mnras/stv1551

Johansson, K. H. (2000), "The quadruple-tank process: a multivariable laboratory process with an adjustable zero". *IEEE Trans. Control Syst.Technol*., vol.8, no.3, pp.456- 465, May. 2000. DOI:10.1109/87.845876

Junshui. M. & Perkins, S. (2003). "Time-series Novelty Detection Using One-class Support Vector Machines". *Proceedings of the International Joint Conference on Neural Networks, 2003,* August-26, Portland, OR, USA, 10.1109/IJCNN.2003.1223670

Liu, F.T., Ting, K. M. & Zhou, Z.H. (2012). "Isolation-Based Anomaly Detection". *ACM Transactions on Knowledge Discovery from Data,* 6(1):1-39 · March 2012, DOI: 10.1145/2133360.2133363

Malhotra, P., Vig, L., Shroff, G. & Agarwal, P. (2015). "Long Short Term Memory Networks for Anomaly Detection in Time Series". *ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8.

Martin, T. E. (1995), *Process control, McGraw-Hill*, New York

Mujica, L., Rodellar, J., Fernandez, A., & Guemes, A. (2010). "Q-statistic and T2-statistic PCA-based measures for damage assessment in structures". *Structural Health Monitoring*, Vol 10, Issue 5, 2011, 10(5) 539–553, DOI: 10.1177/1475921710388972

Ozkan, S., Kara, T., & Arici, M. (2017). "Modelling, simulation and control of quadruple tank process". *10th International Conference on Electrical and Electronics Engineering (ELECO)*. 17523601, Bursa, Turkey, IEEE.

Qin, S. J. (2009). "Data-driven Fault Detection and Diagnosis for Complex Industrial Processes". *Proceedings of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*. Barcelona, Spain, June 30 - July 3, 2009. DOI: 10.3182/20090630-4-ES-2003.00184

Qin, S.J., (2014). "Process Data Analytics in the Era of Big Data". *AIChE Journal*, 60 (9), 3092 – 3100

Rousseeuw, P. J. & Driessen, K. V. (1999), *Technometrics 41*, no. 3 (1999): 212-23. doi:10.2307/1270566

Sakurada. M. & Yairi, T. (2014). "Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction". *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis,* Pages 4, Gold Coast, Australia QLD, Australia — December 02 - 02, 2014, doi:10.1145/2689746.2689747

Samanazari, M., Chaibaksh, A. & Rajabi, S. (2015). "Using One-Class Support Vector Machine for the Fault Diagnosis of an Industrial Once-Through Benson Boiler". *Modares Journal of Electrical Engineering, vol.12, no.3, fall 2015.*

Tagawa, T., Tadokoro, Y. & Yairi, T. (2015). "Structured Denoising Autoencoder for Fault Detection and Analysis". *Proceedings of the Sixth Asian Conference on Machine Learning*, PMLR 39:96-111, 2015.

Tien, D.X., Lim, K. & Jun, L. (2004). "Comparative study of PCA approaches in Process Monitoring and Fault Detection". *30th Annual Conference of the IEEE Industrial Electronics Society*, Busan, Korea, 2-6 November 2004

Yan, W. & Yu, L. (2015). "On accurate and reliable anomaly detection for gas turbine combustors: A deep learning approach". *In Proceedings of the Annual Conference of the Prognostics and Health Management Soci*ety, Coronado, CA, USA, 18–24 October 2015

Yoon, S. & MacGregor, J. F. (2001). "Fault diagnosis with multivariate statistical models part I: using steady state fault signatures". *Journal of Process Control* 11.4 (2001):387-400.), DOI: 10.1016/S0959-1524(00)00008-1

Zope, K. B., Nistala S. H. & Runkana, V. (2018). "Method and System for Anomaly Detection and Diagnosis in Industrial Processes and Equipment", *Indian Patent Filed*, 201921012444, 29 Nov 2019