








A hybrid one-class approach for detecting anomalies in industrial systems

Francisco Zayas-Gato¹  | Esteban Jove¹  | José-Luis Casteleiro-Roca¹  |
 Héctor Quintián¹  | Andrés Piñón-Pazos¹  | Dragan Simić²  |
 José Luis Calvo-Rolle¹ 

¹CTC, Department of Industrial Engineering, CITIC, University of A Coruña, A Coruña, Spain

²Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia

Correspondence

Esteban Jove, CTC, Department of Industrial Engineering, CITIC, University of A Coruña, Avda. 19 de febrero s/n, 15405, Ferrol, A Coruña, Spain.
 Email: esteban.jove@udc.es

Funding information

Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC); Secretaría Xeral de Universidades; European Regional Development Fund; Xunta de Galicia

Abstract

The significant advance of Internet of Things in industrial environments has provided the possibility of monitoring the different variables that come into play in an industrial process. This circumstance allows the supervision of the current state of an industrial plant and the consequent decision making possibilities. Then, the use of anomaly detection techniques are presented as a powerful tool to determine unexpected situations. The present research is based on the implementation of one-class classifiers to detect anomalies in two industrial systems. The proposal is validated using two real datasets registered during different operating points of two industrial plants. To ensure a better performance, a clustering process is developed prior the classifier implementation. Then, local classifiers are trained over each cluster, leading to successful results when they are tested with both real and artificial anomalies. Validation results present in all cases, AUC values above 90%.

KEYWORDS

anomaly detection, clustering, industrial system, one-class, optimization

1 | INTRODUCTION

During last decades of twentieth century and the first decades of twenty first century, the industrial processes have been subjected to a strong development. This has been achieved because of two main different reasons. On the one hand, there has been some technological breakthroughs in terms of instrumentation, optimization, artificial intelligence, among others (Hobday, 1998). On the other hand, it is important to remark the appearance of Internet of Things (IoT). Although it is used in many other areas, such as energy or home automation, IoT impacted significantly the industrial processes. In this context, the term smart industry implicates the development of connected production and intelligent systems have played a significant role (Wortmann & Flüchter, 2015).

Under these circumstances, an early detection of any kind of anomalous situation should be addressed. Then, the smart industries capabilities, are presented as a real interesting way to supervise the correct behaviour of a plant (Wortmann & Flüchter, 2015). To face this task, the combination of these monitoring tools and machine learning techniques is focusing the attention of scientific community (Jove et al., 2019).

Correction added on 4 April 2022 after first online publication: Acknowledgement section has been updated in this version.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Expert Systems* published by John Wiley & Sons Ltd.

To tackle the anomaly detection problem, it is important to start with the definition of what is considered an anomaly. According to Chalapathy et al. (2018) and Chandola et al. (2009), all events that do not have the expected behaviour are considered as anomalous. This concept, which is really general, can be applied over a wide range of fields, such as medicine diagnosis, fraud detection, cybersecurity and, of course, industry (Huang et al., 2012; Kou et al., 2004). This last field is the focus of this research, since the source of the abnormal situation can be diverse, from fault detection in actuators and instruments, to structural damage detection or wrong network performance (Gomes et al., 2019).

Despite the simplicity of the definition presented before, the anomaly detection problem must face several important issues (Chandola et al., 2009). First, it is worth mentioning that, in many applications, it is not possible to know in advance the real behaviour of an anomalous situation before it happens. To overcome this concern, the use of one-class classifiers is a common approach (Tax, 2001). This consists of the use of different techniques to model the dataset behaviour, which is commonly comprised of instances registered during the normal operation. Then, test instances that differ from the implemented model are considered anomalies, since they do not belong to the so-called target class (Rodríguez-Ruiz et al., 2020; Tax, 2001). A critical issue related with the definition of target set modelling is the selection of a decision boundary between what is considered anomalous and what it is not. Another typical issue that must be tackled is the presence of noise in the initial dataset. In this case, instances that might be considered as normal, could present noise, especially in industrial environments (Jove et al., 2019).

Although the one-class approach can deal with the problems detailed before, there are several circumstances where one-class techniques decrease its performance. In many systems, especially in industrial processes, the target set, which is known as the points that belong to the correct operation, can be scattered in different groups. This situation is the result of considering different working points and it is specially concerning when the one-class classifier is implemented using boundary methods to model the target set shape. In these cases, the appearance of anomalies between the different target sub-classes can lead to misclassification. To illustrate the motivation of the work, an example of this undesirable situation is depicted in Figures 1 and 2, where the green dots belong to the target class and the yellow triangle is an anomaly. The consideration of a unique class, presented by a blue boundary, would lead to consider the anomaly as target class. Otherwise, selecting a proper number of sub-classes and applying one-class over each one, would avoid the misclassification situation. It is important to remark that these figures represent the need of finding data groups in the training set, taking into consideration that they are not a description of all one-class classifiers operation.

This issue, added to the idea of contributing to the effective application of artificial intelligence to perform anomaly detection in the smart industry, is the main motivation for this work. Hence, the present research work proposes the combination of clustering algorithms and one-class techniques to carry out the anomaly detection in two industrial plants with a wide range of operating points.

To validate the proposal, two real applications are considered. The first one consists of an industrial system whose main goal is to control the liquid level in a tank. As the idea is to take into account different working points, several tank levels are considered as correct plant behaviour.

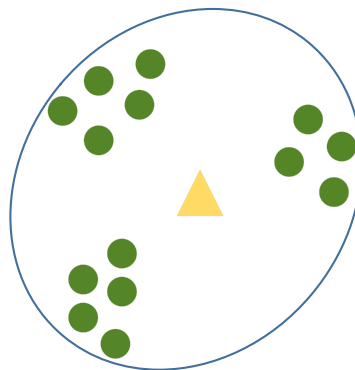


FIGURE 1 Anomaly detection using a one-class approach

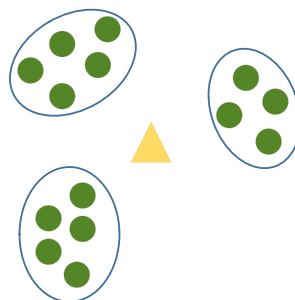


FIGURE 2 Anomaly detection using sub-classes

The second plant has the purpose of manufacturing wind generator blades made of carbon fibre material, obtained by mixing a resin and a catalyst. In this case, the different flows needed to make the blade results in different operating points.

Eight one-class techniques have been applied to model the normal operation of both plants, but, prior to the classification phase, a clustering procedure is followed. The classifier is tested using real created anomalies in the first set, generated by sudden changes in the valve operating status. In the second set, an anomaly generation procedure is followed.

As the registered data can only reflect normal operation of the process, one-class techniques are suitable to accomplish anomaly classification (Tax, 2001). Then, the initial datasets are structured in normal and abnormal sets, taking into consideration only the normal data to carry out the classifier implementation.

The used clustering technique is DBSCAN, which is combined with eight one-class techniques: NCBoP, Autoencoder, Gauss, K-Centers, Minimum Spanning Trees, Parzen Density Estimator, Principal Component Analysis and Support Vector Data Description. The proposal is trained, tested and validated achieving successful results in the anomaly detection task.

This research work is structured according to next sections: Section 2 explains the case of study, Section 3 details the proposal to carry out the anomaly detection. Then, Section 4 describes the techniques used to perform clustering and classification. Sections 5 and 6 details the experiments and results respectively and, finally, Section 7 exposes the conclusions and future works.

2 | CASE OF STUDY

The present section describes the industrial plant used to evaluate the performance of the proposal and the dataset used to implement and validate the classifier proposal.

2.1 | Control level plant

2.1.1 | General description

As stated in the introduction section, this research proposes a one-class anomaly detection classifier combined with a clustering technique. A three dimensional replication of the first plant where the proposal is applied is depicted in Figure 3. The goal of the plant is to control the water level of an objective tank by means of a feeding pipe where the liquid is boosted from a lower tank. The actuator group consists of a three phase motor coupled to a rotatory pump SACI K5T and powered by a Variable Frequency Driver Altivar312–Schneider. Furthermore, the objective tank presents two different discharge valves, one is manual and the other one is an proportional, that will be used to generate anomalies. All pipes have a 1 in diameter.

The instrumentation used to measure the tank level is done using an ultrasonic sensor Banner™ S18UUA, whose working principle consists of sending an ultrasonic wave to the liquid surface and, when it is reflected, it goes back to the sensor again. Then, distance can be calculated by measuring the time elapsed.

As shown in Figure 4, the control loop is implemented using MATLAB software, where the process value is the level percentage of the objective tank and the control signal represents the pump speed. To tackle the nonlinearity of the plant, it is identified on-line as a second order

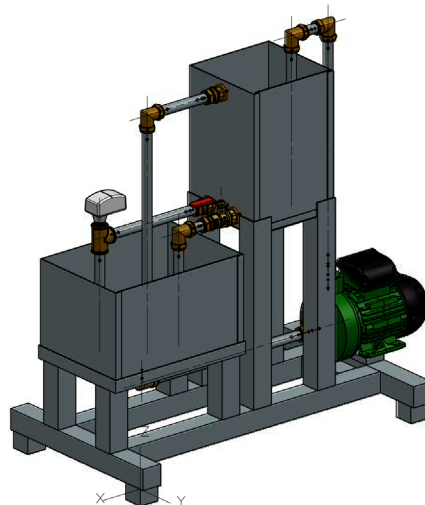


FIGURE 3 Scheme of the control level plant

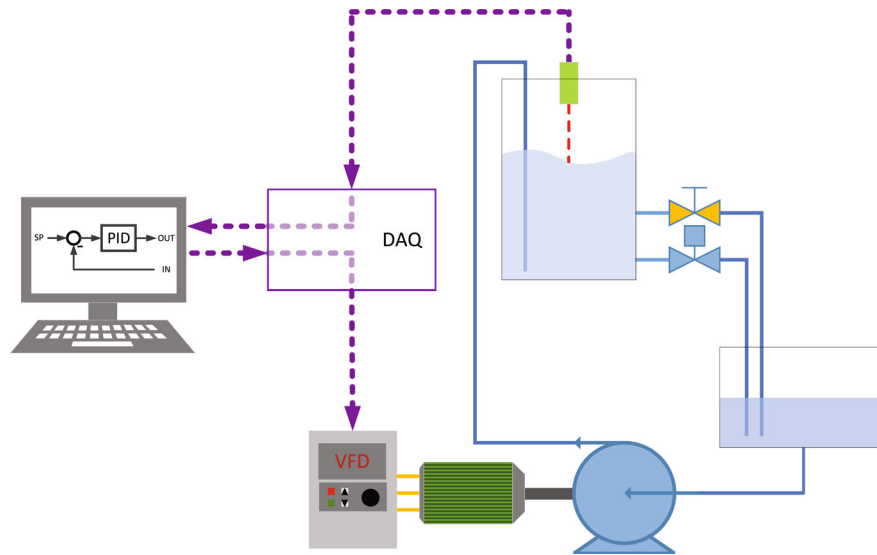


FIGURE 4 Conceptual representation of the control loop

transfer function by means of the Recursive Least Squares (RLS) algorithm, following the Equation (1) (Calvo-Rolle et al., 2014), where b_0 is the open loop gain, k is the system delay and a_1, a_2 are the first and second order coefficient, respectively.

$$P_{\text{system}}(z^{-1}) = \frac{b_0 \cdot z^{-k}}{1 - a_1 \cdot z^{-1} - a_2 \cdot z^{-2}} \quad (1)$$

After this calculation, an adaptive PID is tuned to face nonlinear behaviour according to the Equation (2).

$$P_{\text{PID}}(z^{-1}) = \frac{c_0 + c_1 \cdot z^{-1} + c_2 \cdot z^{-2}}{1 - z^{-1}} \quad (2)$$

where

$$c_0 = \frac{1}{b_0 \cdot T_c^2 \cdot (2 \cdot K_c + 1)}$$

$$c_1 = -a_1 \cdot c_0$$

$$c_2 = -a_2 \cdot c_0$$

Taking into consideration that:

- T_c is the critical period
- K_c is the critical gain

The values measured by the sensor and the control values sent to the Variable Frequency Drive (VFD) that regulates the pump speed are managed by a National Instruments (model USB-6008 12-bit 10 kS/s Multifunction I/O) data acquisition card with a sampling rate of 1 Hz.

2.1.2 | Description of dataset

The dataset used for validating the proposal has been divided in what is considered as correct operation and anomalous situations. An example of the different sets taking into account and how are they labelled is shown in Figure 5. In this case, three different tank levels with three output valve configuration are considered as correct operation. Hence, the target class presents several working points, which is one of the main weaknesses of some one-class techniques. On the other hand, six other plant operations are considered as anomalous. These consists of opening and closing the electric draining valve. Each set is comprised of 5400 instances, for a total number of 16,200 target class samples and 32,400 anomalies.

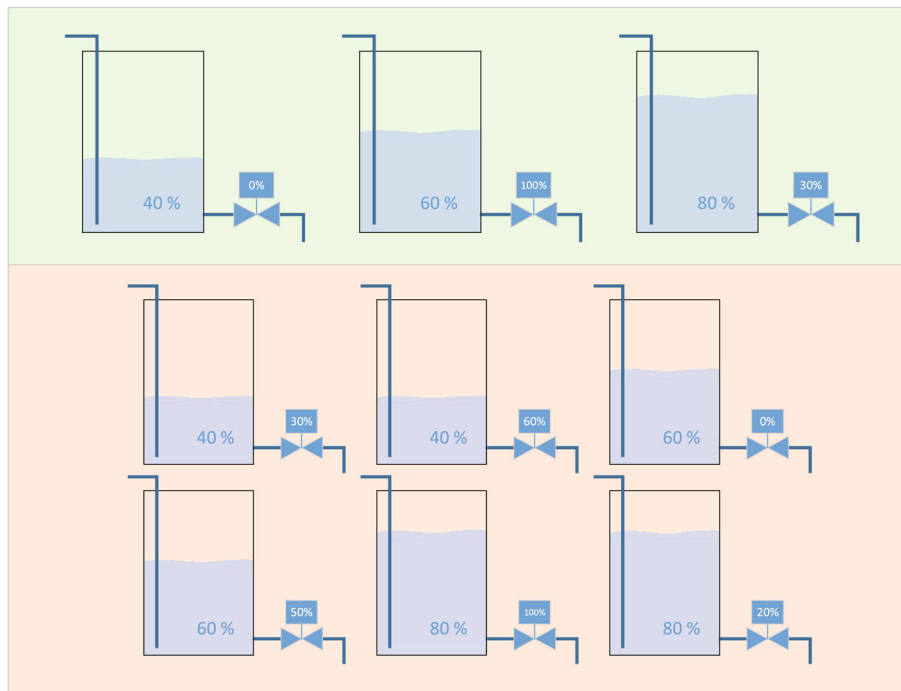


FIGURE 5 Dataset distribution of target class and anomalies classes

Furthermore, the five different features are taken into considered, recorded with sample rate of 1 Hz, contain information about the tank level measured, the PID coefficients c_0 , c_1 and c_2 , and the control signal. Then, each one of these five variables would be configured as classifier inputs to determine the state of the plant.

2.2 | Wind turbine plant

2.2.1 | General description

During fabrication of a turbine blade, the shape of the blade is created by combining a resin with reinforcement of fibre glass or carbon. In order to produce such resin, it is necessary an industrial machine to mix two primary fluids, the epoxy and the catalytic to obtain a final product which presents high tensile and compressive strengths and great chemical resistance (Bank et al., 2018; Mishnaevsky et al., 2017). This industrial mixing machine is schematized in Figure 6.

As can be seen in Figure 6, the system is monitored by means of nine sensors. Each fluid line has one pressure sensor at the output of the pumps (P_{E1} , P_{C1}) and another at the input of the mixing valve (P_{E2} , P_{C2}) and also one flow meter per line (F_{E1} , F_{C1}). Each pump has a speed sensor (S_E , S_C) and finally the flow of mixed material at the output line is also measured (F_M).

The workflow is as follows:

1. Both fluids are stored in separated tanks and sent to the mixing valve by centrifugal pumps controlled through three phase variable frequency drives (VFD).
2. A control system adjusts the pumps speed, acting over each VFD, based on information sampled in the three flow meters to keep a constant output flow ratio and homogeneous material

2.2.2 | Description of dataset

With the aim of generating an accurate anomaly detection model, the industrial plant is monitored by using the nine sensors presented in previous subsection. The control loop has a sampling frequency of 2 Hz, working during 73 min. After removing null measures, the final dataset consists of 8540 samples with nine variables.

To visualize how variables evolve in time, four of the measures during 8 min are presented in Figure 7.

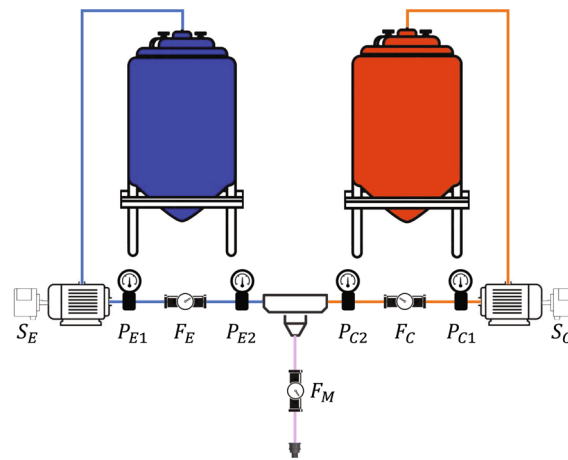


FIGURE 6 Bicomponent mixing machine scheme

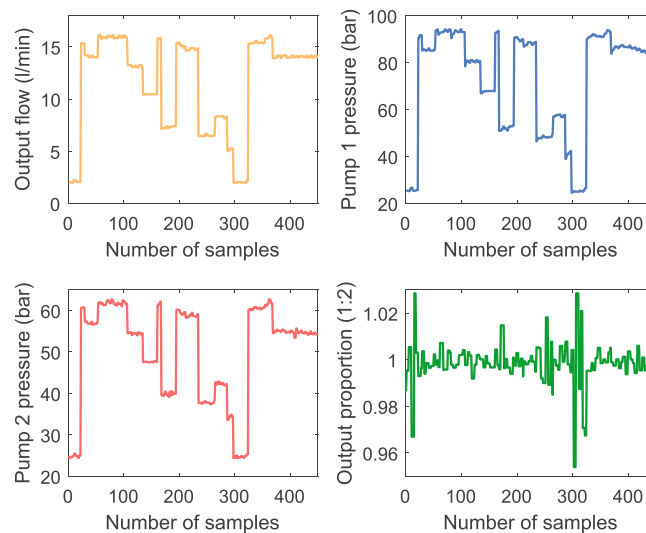


FIGURE 7 Examples of variables measured

Due to the fact that all data samples represent correct working points, the faults have been obtained according to the procedure shown in Figure 8, where j anomalies were generated by modifying a $p\%$ one of the random variables in a total of j random samples from a $M \times N$ dimensional dataset.

From the initial dataset, 1708 samples (20%) of the instances are subjected to the anomaly generation method, modifying a $\pm 15\%$ one of the nine original variables.

3 | CLASSIFIER PROPOSAL

As it was mentioned in the introduction section, one-class techniques are defined following the idea of grouping the data belonging to a normal operation, also known as target set. However, this definition may result in wrong performance, especially when the geometric boundaries are used to establish the limits of the target class. A simple example using a boundary one-class technique in two dimensions is shown in Figure 9. If any anomaly lies in between those clusters it would be classified as a target object, leading to a misclassification. Hence, it would be desirable to divide the data before applying one-class technique over each group, as depicted in Figure 10.

From this general idea, the classifier training is separated in two different steps, represented in Figure 11.

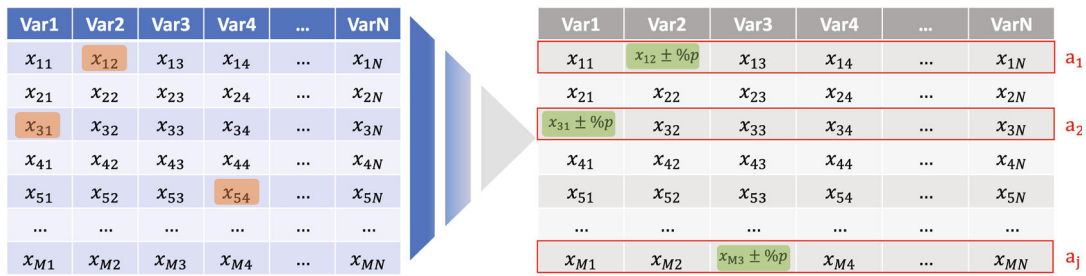


FIGURE 8 Generation of anomalous samples

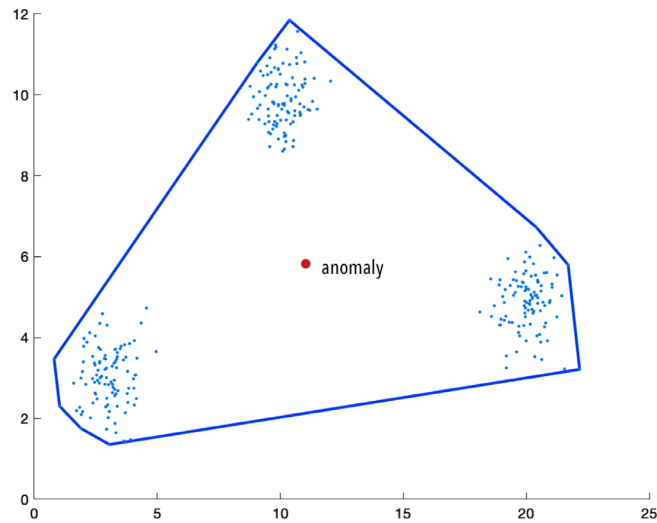


FIGURE 9 One-class boundary concept

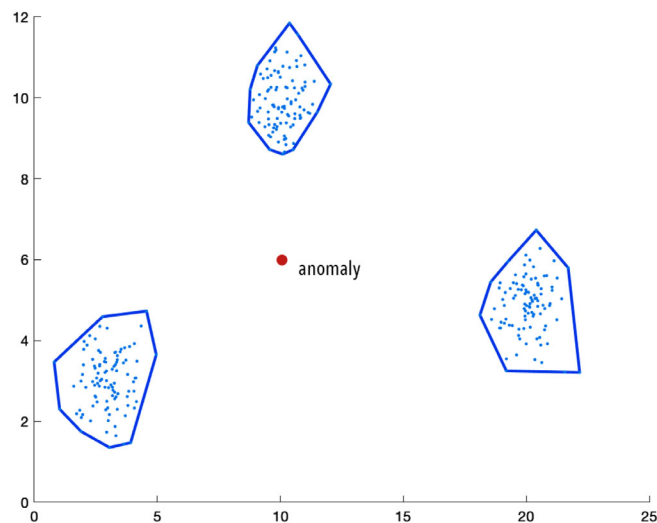


FIGURE 10 One-class boundary concept combined with clustering

1. First, the dataset is divided into N groups without prior knowledge of the proper number of clusters. This research proposes the use of Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm.
2. Then, each of the N groups is treated as independent, so eight different one-class techniques are applied over each cluster of the training set to implement the classifier.

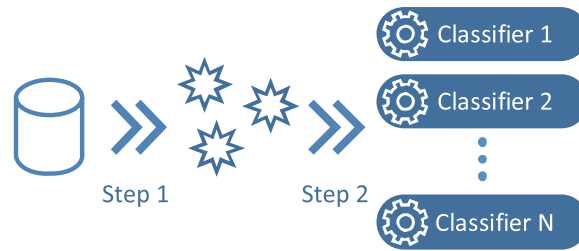


FIGURE 11 Training steps

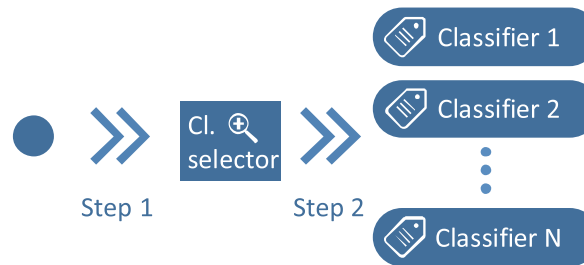


FIGURE 12 Test steps

Once the training is finished and N one-classifiers are implemented, the test phase follows also two steps depicted in Figure 12.

1. In the first step, the test instance is assigned to one of the N clusters determined in the training phase.
2. In a second step, the test instance is sent to the corresponding classifier to label it and determine if it is anomalous.

4 | TECHNIQUES TO ACHIEVE THE CLASSIFIER

The used techniques to implemented the approach detailed in previous section are described in this section.

4.1 | Clustering technique DBSCAN

The density-based spatial clustering of applications with noise (DBSCAN) algorithm is considered the first clustering algorithm based on density (Khan et al., 2014). Besides the proved good performance with many datasets, it received the SIGKDD test-of-time award (Schubert et al., 2017), which gives an idea of its goodness.

This well-known clustering technique was designed to work with data with arbitrary shapes that may present noise (Khan et al., 2014). Its main idea is that for every instance of a cluster, the neighbourhood of a given radius ϵ must contain a minimum amount of instances $MinPnts$. This implies that the neighbourhood cardinality should be above a threshold (Khan et al., 2014). This neighbourhood N_ϵ , of two points a and b is determined by Equation (3).

$$N_\epsilon = \left\{ p \in D / \|b - a\| < \epsilon \right\} \quad (3)$$

where D is the database of instances. To consider a point p as core point, it must have at least $MinPnts$ points inside its neighbours, as shown in Equation (4). It is considered as non-core otherwise (Khan et al., 2014).

$$N_\epsilon(p) > MinPnts \quad (4)$$

A new cluster is created each time a new point satisfies this condition of having at least $MinPnts$ neighbourhood. From this core points, more points are searched until no more object can be included in the cluster (Khan et al., 2014).

The DBSCAN clustering technique presents the main advantage of discovering the groups with arbitrary shapes, both linear and nonlinear. It also has the advantage of not having to determine the number of clusters in which the dataset is divided. Furthermore, its good performance use has been widely used especially with large datasets (Birant & Kut, 2007).

4.2 | One-class techniques

4.2.1 | Non convex boundary over projections

The Non Convex Boundary over Projections (NCBoP) algorithm is based on the idea of modelling the shape of the target set by means of the calculation of non-convex hull (Jove et al., 2021). This novel one-class classification algorithm overcomes the weaknesses presented by well-known Convex Hull over random projections (Casale et al., 2014). The main basis of this method is to approximate the boundaries of a dataset $S \in \mathbb{R}^n$ using the non-convex hull over a π random projections on 2D planes and then, determine the non-convex limits on that plane, reducing in this way the complexity of calculating the non-convex limits over \mathbb{R}^n .

The NCBoP calculates one non-convex polygon for each π 2D random projections, by selecting an starting point as the lowest y coordinate among all points project over the p_i plane. Then it is calculated its k-nearest points and they are ordered based on the polar angle to finally kept only the further one from the starting point. Once this first couple of points are calculated, a stack structure is created including the next third point. Then it is checked if the next point in the list turns left (pushed into the stack) or right (point on the top is removed). This process is repeated until it is arrived again to the starting point. When the training process finish, all the points are in the non-convex polygon created by the algorithm. Once the training phase is over, the criteria to determine if a new test point is anomalous is the following: if the data is out of at least one of the π projected hulls, the data is considered as anomalous.

4.2.2 | Autoencoder

This Autoencoder technique is based on the use of an Artificial Neural Network (ANN) to determine the appearance of anomalies. The most common ANN configuration is structured in one input layer, one or more hidden layer and an output layer, whose neurons are connected with wheighted links (Sakurada & Yairi, 2014). The main idea of this technique is to reconstruct the input p at the output \hat{p} by means of a nonlinear dimensional reduction in the hidden layer. Hence, the number of neurons in the hidden layer are less than the number of inputs.

Once the network is trained using only information from the training set, that data with different behaviour is supposed to present significant different in the hidden layer subspace. This implies that a test data q should have great reconstruction error, which is defined as $\|q - \hat{q}\|$. This value is the criteria to label the data as anomalous (Tax, 2001).

4.2.3 | Gaussian model

A especial approach to face anomaly detection using one-class techniques is based on the idea of using density functions. Although more complex topologies considering Gaussian use (Oza & Patel, 2018), the most direct way to follow this approach consists of applying a normal or Gaussian distribution function over the target set (Tax, 2001). This function is calculated from the target set, which is the same as the training set.

Once the mean vector and covariance matrix are known, the criteria to determine the anomalous nature of a test sample is based on the value obtained of this instance in the Gaussian function. The simplicity of this idea has the corresponding low computational cost, which is a significant advantage in large datasets with normal shapes.

A simplified example of a Gaussian function (blue line) with a one-dimension set is depicted in Figure 13, where the criteria to determine anomalous nature of a test instance is determined by the threshold (red lines). This value is established during the training phase.

4.2.4 | K-centers

The K-centers method aims to obtain the model of the target set by covering all the points with K hyper-spheres, all of them with equal radius. These hyper-spheres are set looking for a minimisation of the maximum distances between each instance and its center (Tax, 2001). This process can be divided into two steps. First, each training instance is assigned to the cluster whose center is the closest. These centers are randomly selected. Then, the position of each center is adjusted to ensure the minimum distance between it and the rest of points of this cluster. To avoid local minimums, these two steps are repeated several times and the configuration with the optimum solution is chosen (Japkowicz, 1999).

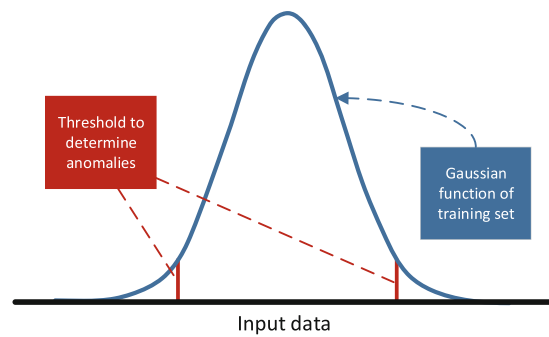


FIGURE 13 Example of Gaussian model for one-class in \mathbb{R}^1

When the final K-Centers configuration is chosen, a new test instance would be labelled as target class data if it is inside of one of the hyperspheres implemented during the training phase (Japkowicz, 1999).

4.2.5 | Minimum spanning tree

This one-class method is based on target set modelling by means of the structure obtained by a Minimum Spanning Tree (MST). It relies on the assumption that two points $p_i, p_j \in \mathbb{R}^n$ that belong to target class should be neighbours in \mathbb{R}^n representation (Juszczak et al., 2009). Then, a linear transformation can be found for these points and all points considered as target class. As this set commonly contains more than two objects, then, more than one transformation can be considered. Hence, for a dataset D with n instances, $(n - 1)$ linear transformations can be found (Juszczak et al., 2009).

Then, the MST consists of a set of edges e_{ij} that specify the linear transformations of each point. It is implemented a graph ensuring the absence of loops and minimizing the total length of the edges. Once the MST is trained with data from the target set, the distance between a test object to its nearest edge projection is the criteria to determine an anomalous situation (Juszczak et al., 2009).

4.2.6 | Parzen density estimator

The idea previously mentioned in the Gaussian model can be applied when it is used the non-parametric Parzen density estimator (PDE) (Parzen, 1962), whose performance has proven to be successful in many UCI repositories (Casale et al., 2014). Although it is based on the idea of using density function, it presents the advantage of having good performance even when the data do not have Gaussian shape (Mazhelis, 2006).

This method can be considered a mixture of Gaussian functions centered on the individual training instances, with diagonal covariance matrices $\sum_i = h \cdot I$, where the optimal width of the kernel is adjusted during the training set, using a maximum likelihood solution (Tax, 2001). This technique shows better results the size is greater (Cohen et al., 2008).

4.2.7 | Principal component analysis

The Principal Component Analysis (PCA) technique has been commonly used for dimensional reduction problems (Chiang et al., 2000; Wu & Zhang, 2001). It is based on the concept of calculating the Principal Components of the training set, which represents the directions where the data have greater variability. This vectors are computed from the eigenvalues of the covariance matrix, which is follows a relatively simple calculation. Then, using this principal components, the dataset can be projected onto a subspace with lower dimension. This dimensional reduction process can be exploited as it happens with Autoencoder configuration by means of the reconstruction error, which is computed as the distance between the original and the projected data.

Hence, the number of eigenvectors, known as components, can be at most the same as the number of variables. The criteria followed to decide if a test instance belongs to the target class is based on the reconstruction error. This is computed as the difference between the original point and the point projected in the subspace. This technique offers good results when the subspace is clearly linear (Tax, 2001).

Figure 14 represents an example of how the first component (dashed lines) of a training set in \mathbb{R}^2 (blue dots) can be used to detect anomalies. When a new test sample (red dot) presents greater distance to the first component projection than the distances of training set, it is considered as anomalous.

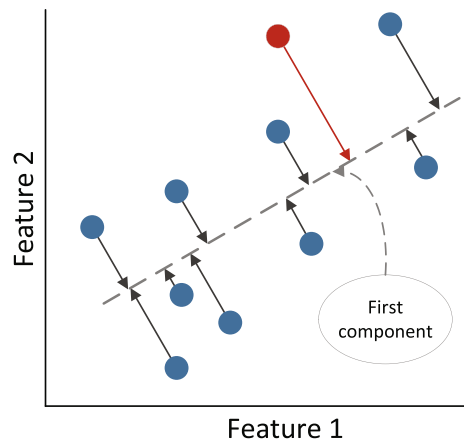


FIGURE 14 Example of PCA for one-class tasks in \mathbb{R}^2

4.2.8 | Support vector data description

The Support Vector Data Description (SVDD) technique was created in (Tax, 2001) as a boundary one class technique. This method is derived from well-known Support Vector Machine (SVM) (Miao et al., 2018), which is a supervised algorithm whose main procedure consists of mapping the training onto a high dimensional feature space by means a kernel function and then, a hyper-plane is constructed to perform classification (Rebentrost et al., 2014).

The SVDD technique presents an analogous approach but, in this case, instead a hyper-plane, a hyper-sphere is implemented to delimit the target class shape (Sanchez-Hernandez et al., 2007). The radius and the centre are tuned during the training phase and then, the anomalous point should be placed outside the implemented hyper-volume. Although this technique presents successful performance over a wide amount of sets and applications, it presents great computation effort compared with most of the previous detailed techniques.

5 | EXPERIMENTS

As the main proposal of this research is to improve the one-class classifiers performance over systems with different operating points, a comparative analysis of eight different one-class techniques is carried out combined with an clustering stage. Furthermore, the achieved classifiers are compared with the ones obtained without the prior clustering stage. The different hyper-parameters are tested for each classification technique are shown in Table 1:

Besides all the hyper-parameters tested for each technique, the data is considered with three different configurations:

- Type 0: Data without pre-processing.
- Type 1: Each variable is normalized using a 0 to 1 scaling range (see Equation (5), where min and max are the minimum and maximum values registered in a variable).
- Type 2: Each variable is normalized using a Z-Score conversion (see Equation (6), where μ is the mean and σ is the standard deviation of a variable).

$$p_{iNorm} = \frac{p_i - \min}{\max - \min} \quad (5)$$

$$p_{iNorm} = \frac{p_i - \mu}{\sigma} \quad (6)$$

- Initially, with 90% of the data, a k -fold cross validation method is followed with $k = 5$, as depicted in Figure 15. This procedure consists of dividing in a random way the target set in five different groups. Then, five classifiers are implemented using the 80% of the data and leaving the 20% to the test phase. The use five folds ensures that all instances are considered for the training and test phases. Then, the classifier is tested using the non-target set and 20% of the target data.

TABLE 1 Tested configuration for each technique

Technique	Hyperparameter 1	Values tested	Hyperparameter 2	Values tested
NCBoP	Number of projections	20, 50, 200, 500	Expansion factor	0,5 to 1,5 with step 0,1
Autoenc	Neurons in the hidden layer	1, 2, 3, 4	Training rejection rate	0, 5, 10, 15, 20, 25
Gauss	Training rejection rate	0, 5, 10, 15, 20, 25	-	-
K-Centers	Number of Centers	1 to 15 with step 1	Training rejection rate	0, 5, 10, 15, 20, 25
MST	Training rejection rate	0, 5, 10, 15, 20, 25	-	-
Parzen	Training rejection rate	0, 5, 10, 15, 20, 25	-	-
PCA	Number of components	1, 2, 3, 4	Training rejection rate	0, 5, 10, 15, 20, 25
SVDD	Kernel Width	0,5 to 10 with step 0,5	Training rejection rate	0, 5, 10, 15, 20, 25

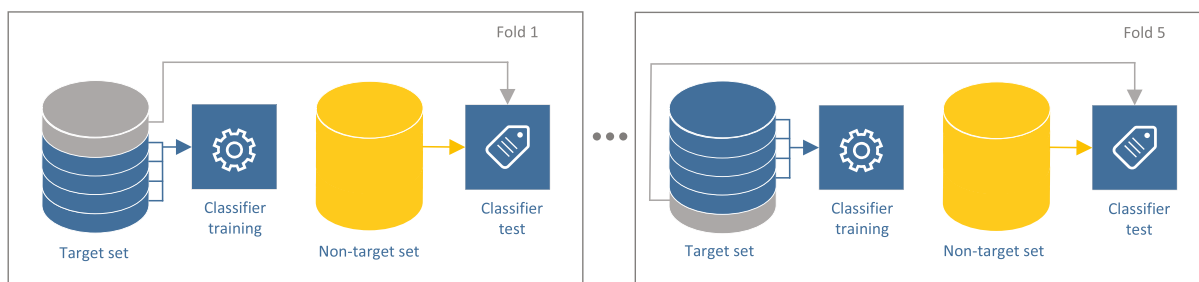


FIGURE 15 Schematic description of k -fold with $k=5$

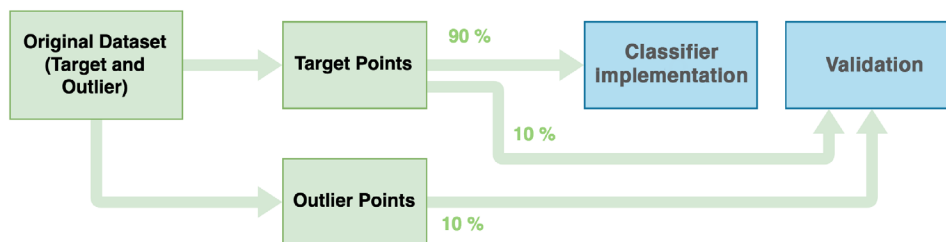


FIGURE 16 Final values of the validation process

- Once the best configuration for each cluster is known, the classifiers are trained with the 90% left in the previous phase. The performance of each technique is validated using the 10% of the data that were not used to train or test the classifier. This procedure is depicted in Figure 16.

The performance of each classifier configuration is evaluated using the well-known Area Under the Receiving Operating Characteristics Curve (AUC) measure (Fawcett, 2006). The AUC combines the true positive and false positive rates, obtaining a unique measure of the classifier performance. From a statistical point of view, this value represents the probability of classifying a random positive instance as positive (Fawcett, 2006). Furthermore, in contrast to other measures like sensitivity, precision or recall, AUC is not sensitive to class distribution, which is a significant advantage especially in one-class tasks (Bradley, 1997).

6 | RESULTS

6.1 | Control level plant

An initial approach without a clustering process is followed to evaluate the importance of data grouping. The results of this experiments are presented in Table 2. Then a clustering stage is applied using DBSCAN, resulting in three different clusters. Hence, the highest AUC achieved for each technique and its corresponding configuration for cluster 1, 2 and 3 are represented in Tables 3–5.

TABLE 2 Performance without clustering for control level plant set. Global classifier

Technique	Type	Hyp. 1	Hyp. 2	AUC(%)	t_i (s)	t_L (μ s)
APEB	0	1	1000	76.86	217.91	377485.44
Autoenc	1	10	15	81.95	2.15	0.94
Gauss	1	5	-	73.91	0.01	0.27
K-Centers	1	2	15	84.41	233.92	2398.45
MST	2	0	-	94.65	232.12	1995.94
Parzen	1	5	-	87.77	118.85	210.14
PCA	1	1	15	72.47	0.06	0.43
SVDD	0	2	0	85.53	2043.548	8.32

TABLE 3 Performance for control level plant set. Classifier of cluster 1

Technique	Type	Hyp. 1	Hyp. 2	AUC(%)	t_i (s)	t_L (μ s)
NCBoP	0	500	1.2	99.57	31.84	47574.62
Autoenc	2	1	10	92.12	0.38	0.82
Gauss	1	5	-	94.01	0.01	0.34
K-Centers	1	7	0	99.51	55.95	42.67
MST	1	0	-	99.36	4.44	632.06
Parzen	2	10	-	94.56	5.47	54.41
PCA	0	1	10	94.42	0.05	0.38
SVDD	2	1	0	99.52	96.61	0.98

TABLE 4 Performance for control level plant set. Classifier of cluster 2

Technique	Type	Hyp. 1	Hyp. 2	AUC(%)	t_i (s)	t_L (μ s)
NCBoP	0	500	1.00		33.13	49463.15
Autoenc	1	1	10	94.14	0.26	0.83
Gauss	0	0	-	97.43	0.01	0.19
K-Centers	1	3	5	95.14	23.28	47.33
MST	1	5	-	99.05	4.53	630.87
Parzen	1	10	-	94.82	8.51	50.88
PCA	2	4	10	94.98	0.02	0.20
SVDD	2	1	0	93.86	73.88	0.66

TABLE 5 Performance for control level plant set. Classifier of cluster 3

Technique	Type	Hyp. 1	Hyp. 2	AUC(%)	t_i (s)	t_L (μ s)
NCBoP	0	500	1.2	99.85	36.26	54123.52
Autoenc	0	1	10	95.34	4.05	0.94
Gauss	1	0	-	99.82	0.01	0.33
K-Centers	1	7	0	99.99	69.72	49.04
MST	1	0	-	99.98	4.42	632.70
Parzen	2	10	-	94.64	5.40	46.91
PCA	1	1	0	99.99	0.02	0.20
SVDD	0	2.5	0	99.29	75.35	1.64

First it can be seen how the results achieved without clustering are worse in terms of AUC and training times, confirming that the grouping is suitable for this dataset.

NCBoP, using 500 projections, an expansion parameter of 1.2, presents the best results in terms of classification performance for the first cluster. Although the time needed to label a new test is greater than 47 ms, it is far below the sampling time of the system. Then, this

configuration is accepted for the first cluster. In case of critical sampling times, other techniques such as SVDD, K-Centers or MST could be considered.

In respect of cluster 2, the MST is the algorithm that achieves the greatest performance. This technique is significantly better than the rest, with a difference of at least 1.5% comparing to the rest of techniques.

Finally, in the third cluster, almost all techniques achieve their best results. Although PCA, K-Centers and MST are tied, the technique selected is PCA, since it presents the lowest computational cost and labelling time. This value corresponds to a configuration with a pre-processing stage, one principal component and a training rejection rate of 0.

According to the achieved results, the selected topology is structured using NCBOP for cluster 1, MST for cluster 2 and K-Centers for cluster 3. Then, the data left unused in the training/test phase is used to validate the classifier. The configuration and results for this phase is represented in Table 6.

TABLE 6 Classifier performance according to each cluster. Control level plant set

Cluster	Technique	Type	Hyp. 1	Hyp. 2	AUC (%)
CI 1	NCBoP	0	500	1.2	95.88
CI 2	MST	1	5	-	96.78
CI 3	PCA	1	1	0	96.42

TABLE 7 Performance without clustering for wind turbine set. Global classifier

Technique	Type	Hyp. 1	Hyp. 2	AUC(%)	t_t (s)	t_L (μ s)
NCBoP	0	1	100	84.77	10.35	256096.08
Autoenc	1	16	5	89.27	10.94	5.82
Gauss	1	5	5	84.40	0.01	2.82
K-Centers	0	20	10	69.39	35.27	70.58
MST	0	5	-	88.32	7.99	1477.76
Parzen	1	0	-	89.30	6.37	62.22
PCA	1	4	5	90.49	0.03	1.64
SVDD	2	1	0	86.07	206.00	5.04

TABLE 8 Best AUC value for each cluster. Wind turbine set

Cluster	Technique	Type	Hyp. 1	Hyp. 2	AUC(%)	t_t (s)	t_L (μ s)
1	MST	2	10	-	96.78	1.11	570.25
2	K-Centers	2	7	0	95.59	0.14	2.00
3	MST	1	5	-	89.27	0.04	97.37
4	MST	1	5	-	98.13	0.01	30.43
5	MST	1	15	-	95.26	0.01	29.95
6	MST	2	10	-	92.78	0.02	76.16
7	MST	1	15	-	90.68	0.01	37.44
8	NCBoP	0	50	1.2	94.58	0.20	1591.50
9	K-Centers	2	7	0	95.35	0.21	5.33
10	MST	1	10	-	95.98	0.03	83.91
11	MST	2	15	-	96.31	0.04	93.09
12	NCBoP	0	10	1.2	96.58	0.03	265.95
13	Gauss	1	0	-	94.36	0.01	0.59
14	MST	1	15	-	97.07	0.02	50.44

TABLE 9 Classifier performance (%) in validation phase for wind turbine plant dataset

CI1	CI2	CI3	CI4	CI5	CI6	CI7	CI8	CI9	CI10	CI11	CI12	CI13	CI14
93.90	92.15	87.42	96.25	92.33	89.68	88.02	93.57	91.36	93.28	91.87	94.87	90.13	95.95

6.2 | Wind turbine plant

The results achieved without the use of DBSCAN are presented in Table 7. They are compared with the ones obtained with prior clustering process, which organized the dataset in 14 different groups. Then, taking into account that at each cluster, 18 different techniques are applied, 112 classifiers are implemented. To summarize the results, the technique with greatest AUC for each cluster is represented in Table 8. The best results using DBSCAN always outperform the results achieved by the global classifier. Furthermore, it is important to emphasize that MST outperforms the rest of classifiers in 9 out of 14 clusters. K-Centers and NCBOP achieve the best results in two clusters, and Gauss in one.

Using the selected topology of Table 8, the data unused in the training/test stage is used to validate the classifier. The configuration and results of this validation is shown in Table 9.

7 | CONCLUSIONS AND FUTURE WORKS

This work faces the problem of anomaly detection in two different industrial environments where the systems work at different operating points. The proposal is evaluated using two plants with different features, considering real anomalies and artificial anomalies. The proposal takes advantage of clustering algorithm DBSCAN to divide the target set prior the classifier implementation. This clustering process improves the classifiers results in terms of AUC and also training times. The results obtained reveals that the normal operation is clustered in three different groups for the first plant and 14 clusters for the second one.

Focusing on control level plant, the final topology leads to significant high classification rates, with a validation results of 95.88% in cluster 1, 96.78% in cluster 2 and 96.42 in cluster 3%. Regarding wind turbine plant, the validation results are at least above 90% in all clusters, which is a significant good performance.

This approach can be used to improve the classification when the target set is scattered in different clusters or groups. Using this idea, an early detection of anomalous situations is more feasible, increasing the systems optimisation. This proposal could also help to avoid the propagation of wrong sensor measurements or actuator failures, reducing the corrective maintenance costs and assisting the predictive maintenance schedule.

An interesting research line to continue this work could consist of training the classifier on-line. This would have the strength of training the anomaly detection system during its operations, with the possibility of learning its evolution. However, this idea has the weakness of increasing the target class boundaries, running the risk of assuming that anomalies are normal operation points.

Finally, different anomalies in other plant components, such as sensors and actuators could be considered to implement a distributed topology. Training the classifier with prior knowledge of the anomalies generated would help to identify the source of the wrong performance and isolate this part of the facility.

ACKNOWLEDGEMENTS

CITIC, as a Research Center of the University System of Galicia, is funded by Consellería de Educación, Universidade e Formación Profesional of the Xunta de Galicia through the European Regional Development Fund (ERDF) and the Secretaría Xeral de Universidades (Ref. ED431G 2019/01). Universidade da Coruña/CISUG, for funding the open access charge.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Francisco Zayas-Gato  <https://orcid.org/0000-0002-0994-1961>

Esteban Jove  <https://orcid.org/0000-0002-0625-359X>

José-Luis Casteleiro-Roca  <https://orcid.org/0000-0001-9740-6477>

Héctor Quintián  <https://orcid.org/0000-0002-0268-7999>

Andrés Piñón-Pazos  <https://orcid.org/0000-0002-0116-8361>

Dragan Simić  <https://orcid.org/0000-0001-7392-9767>

José Luis Calvo-Rolle  <https://orcid.org/0000-0002-2333-8405>

REFERENCES

- Bank, L. C., Arias, F. R., Yazdanbakhsh, A., Gentry, T. R., Al-Haddad, T., Chen, J.-F., & Morrow, R. (2018). Concepts for reusing composite materials from decommissioned wind turbine blades in affordable housing. *Recycling*, 3(1), 3.
- Birant, D., & Kut, A. (2007). St-dbscan: An algorithm for clustering spatial-temporal data. *Data & Knowledge Engineering*, 60(1), 208–221.
- Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7), 1145–1159.
- Calvo-Rolle, J. L., Fontenla-Romero, O., Pérez-Sánchez, B., & Guijarro-Berdiñas, B. (2014). Adaptive inverse control using an online learning algorithm for neural networks. *Informatica*, 25(3), 401–414.
- Casale, P., Pujol, O., & Radeva, P. (2014). Approximate polytope ensemble for one-class classification. *Pattern Recognition*, 47(2), 854–864.
- Chalapathy, R., Menon, A. K., & Chawla, S. (2018). Anomaly detection using one-class neural networks. arXiv preprint arXiv:1802.06360.
- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, 41(3), 15.
- Chiang, L. H., Russell, E. L., & Braatz, R. D. (2000). *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media.
- Cohen, G., Sax, H., & Geissbuhler, A. (2008). Novelty detection using one-class parzen density estimator. An application to surveillance of nosocomial infections. *Mie*, 136, 21–26.
- Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, 27(8), 861–874.
- Gomes, I. L. R., Melicio, R., Mendes, V. M. F., & Pousinho, H. M. I. (2019). Wind power with energy storage arbitrage in day-ahead market by a stochastic MILP approach. *Logic Journal of the IGPL*, 28(4), 570–582. <https://doi.org/10.1093/jigpal/jzz054>
- Hobday, M. (1998). Product complexity, innovation and industrial organisation. *Research Policy*, 26(6), 689–710.
- Huang, Z., Lu, X., & Duan, H. (2012). Anomaly detection in clinical processes. *AMIA Annual Symposium proceedings/AMIA Symposium. AMIA Symposium, 2012*, 370–379.
- Japkowicz, N. (1999). Concept-learning in the absence of counter-examples: An autoassociation-based approach to classification.
- Jove, E., Casteleiro-Roca, J.-L., Quintián, H., Méndez-Pérez, J.-A., & Calvo-Rolle, J. L. (2019). Virtual sensor for fault detection, isolation and data recovery for bicomponent mixing machine monitoring. *Informatica*, 30(4), 671–687.
- Jove, E., Casteleiro-Roca, J.-L., Quintián, H., Méndez-Pérez, J.-A., & Calvo-Rolle, J. L. (2021). A new method for anomaly detection based on non-convex boundaries with random two-dimensional projections. *Information Fusion*, 65, 50–57.
- Juszczak, P., Tax, D. M., Pe, E., Duin, R. P., et al. (2009). Minimum spanning tree based one-class classifier. *Neurocomputing*, 72(7–9), 1859–1869.
- Khan, K., Rehman, S. U., Aziz, K., Fong, S., & Sarasvady, S. (2014). Dbscan: Past, present and future. In *The Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2014)*, pp. 232–238.
- Kou, Y., Lu, C.-T., Sirwongwattana, S., & Huang, Y.-P. (2004). Survey of fraud detection techniques. In *IEEE International Conference on Networking, Sensing and Control, 2004*, Vol. 2, pp. 749–754.
- Mazhelis, O. (2006). One-class classifiers: A review and analysis of suitability in the context of mobile-masquerader detection. *South African Computer Journal*, 2006(36), 29–48.
- Miao, X., Liu, Y., Zhao, H., & Li, C. (2018). Distributed online one-class support vector machine for anomaly detection over networks. *IEEE Transactions on Cybernetics*, 49(4), 1475–1488.
- Mishnaevsky, L., Branner, K., Petersen, H. N., Beauson, J., McGugan, M., & Sørensen, B. F. (2017). Materials for wind turbine blades: An overview. *Materials*, 10(11), 1285.
- Oza, P., & Patel, V. M. (2018). One-class convolutional neural network. *IEEE Signal Processing Letters*, 26(2), 277–281.
- Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33(3), 1065–1076.
- Rebentrost, P., Mohseni, M., & Lloyd, S. (2014). Quantum support vector machine for big data classification. *Physical Review Letters*, 113, 130503. <https://doi.org/10.1103/PhysRevLett.113.130503>
- Rodríguez-Ruiz, J., Mata-Sánchez, J. I., Monroy, R., Loyola-González, O., & López-Cuevas, A. (2020). A one-class classification approach for bot detection on twitter. *Computers & Security*, 91, 101715.
- Sakurada, M., & Yairi, T. (2014). Anomaly detection using autoencoders with nonlinear dimensionality reduction. In *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis* (p. 4).
- Sanchez-Hernandez, C., Boyd, D. S., & Foody, G. M. (2007). One-class classification for mapping a specific land-cover class: Svdd classification of fenland. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4), 1061–1073.
- Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). Dbscan revisited, revisited: Why and how you should (still) use dbscan. *ACM Transactions on Database Systems (TODS)*, 42(3), 1–21.
- Tax, D. M. J. (2001). One-class classification: concept-learning in the absence of counter-examples [PhD. thesis]. *Delft University of Technology*.
- Wortmann, F., & Flüchter, K. (2015). Internet of things. *Business & Information Systems Engineering*, 57(3), 221–224.
- Wu, J., & Zhang, X. (2001). A PCA classifier and its application in vehicle detection. In *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (cat. No. 01ch37222)*, Vol. 1, pp. 600–604.

AUTHOR BIOGRAPHIES

Francisco Zayas-Gato received B.S. degree in Industrial Electronics and Automation Engineering and M.S. degree in Efficiency and Energy use from the University of A Coruña (Spain). After 9 years working in automation and electronics in industries such as energy, steel, naval or alimentary, among others, he joined the University of A Coruña, as Professor of Power Electronics and Automation in the Faculty of Engineering since 2020. His main research interests are focused on the application of intelligent techniques for prediction and anomaly detection in industrial systems.

Esteban Jove received M.S. degree in Industrial Engineering from the University of Leon in 2014. After two years working in the automotive industry, he joined the University of A Coruña, Spain, where he is Professor of in the Energy Efficiency Master's Degree since 2016. In 2020,



he received the Ph.D in the Industrial, Environmental and Computer Science engineering program, in the University of La Laguna. His main research interests have been focused on the application of intelligent techniques to model, optimize and detect anomalies in industrial and energetic systems.

José Luis Casteleiro Roca has received M.S. and Ph.D. degrees in Industrial Engineering from the University of Leon in 2004 and 2007, respectively. He is an Associate Professor of Automatic Control and the head of the Industrial Engineering Department, Faculty of Engineering, University of A Coruña, Spain. His main research areas are related with the application of expert system technologies in diagnosis and control systems and in intelligent training systems for control engineering, systems optimization and education.

Héctor Quintián received the MS degrees in Industrial Engineering from the University of Leon, Leon, Spain, in 2010, and his PhD in Computer Science in 2016. He is an associate professor of Automatic Control at University of A Coruña, Spain. His main research interests have been centered in neural networks, unsupervised systems, and intelligent training systems for control engineering, optimization, and education.

Andrés Piñón-Pazos was born in A Coruna, Spain, in 1969. He received the M.S. degrees in Industrial Engineering from the University of Coruna, Coruna, Spain, in 1998. He is Associate Professor in the Systems Engineering and Automatic Control area of Industrial Engineering Department, Faculty of Engineering, University of A Coruna, Spain. His main research is centered on sensors, electronic instrumentation and fieldbus sensors.

Dragan Simić received the Ph.D. degree at the Department of Mathematics and Computer Sciences, Faculty of Sciences, University of Novi Sad, Novi Sad, Serbia. More than 25 years he worked in industry. He is currently a Full Professor at the Department of Transport Engineering, Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia. His current research interest includes: information systems, decision support systems, hybrid artificial intelligence systems, soft computing, and their applications in logistics, production, financial forecasting, and medicine. He is a member of the International Federation for Information Processing (IFIP) – and represents Republic of Serbia in Council and Technical Committee TC 8 – Information Systems from 2007.

José Luis Calvo Rolle received M.S. and Ph.D. degrees in Industrial Engineering from the University of Leon in 2004 and 2007, respectively. He is an Associate Professor of Automatic Control and the head of the Industrial Engineering Department, Faculty of Engineering, University of A Coruña, Spain. His main research areas are associated with the application of expert system.

How to cite this article: Zayas-Gato, F., Jove, E., Casteleiro-Roca, J.-L., Quintián, H., Piñón-Pazos, A., Simić, D., & Calvo-Rolle, J. L. (2022). A hybrid one-class approach for detecting anomalies in industrial systems. *Expert Systems*, 39(9), e12990. <https://doi.org/10.1111/exsy.12990>