Using Computational Intelligence for the Safety Assessment of Oil and Gas Pipelines: A Survey

Abduljalil Mohamed, Mohamed Salah Hamdi and Sofiène Tahar

Abstract The applicability of intelligent techniques for the safety assessment of oil and gas pipelines is investigated in this study. Crude oil and natural gas are usually transmitted through metallic pipelines. Working under unforgiving environments, these pipelines may extend to hundreds of kilometers, which make them very susceptible to physical damage such as dents, cracks, corrosion, etc. These defects, if not managed properly, can lead to catastrophic consequences in terms of both financial losses and human life. Thus, effective and efficient systems for pipeline safety assessment that are capable of detecting defects, estimating defects sizes, and classifying defects are urgently needed. Such systems often require collecting diagnostic data that are gathered using different monitoring tools such as ultrasound, magnetic flux leakage, and Closed Circuit Television (CCTV) surveys. The volume of the data collected by these tools is staggering. Relying on traditional pipeline safety assessment techniques to analyze such huge data is neither efficient nor effective. Intelligent techniques such as data mining techniques, neural networks, and hybrid neuro-fuzzy systems are promising alternatives. In this paper, different intelligent techniques proposed in the literature are examined; and their merits and shortcomings are highlighted.

Keywords Oil and gas pipelines · Safety assessment · Big data · Computational intelligence · Data mining · Artificial neural networks · Hybrid neuro-fuzzy systems · Artificial intelligence · Defect sizing · Magnetic flux leakage

---

A. Mohamed (✉) · M.S. Hamdi
Department of Information Systems, Ahmed Bin Mohamed Military College,
P. O. Box 22713, Doha, Qatar
e-mail: ajamoham@abmmc.edu.qa

M.S. Hamdi
e-mail: mshamdi@abmmc.edu.qa

S. Tahar
Department of Electrical and Computer Engineering, Concordia University,
1515 St. Catherine W, Montreal, Canada
e-mail: tahar@ece.concordia.ca

W. Pedrycz and S.-M. Chen (eds.), *Data Science and Big Data: An Environment of Computational Intelligence*, Studies in Big Data 24,
DOI 10.1007/978-3-319-53474-9_9

1 Introduction

Oil and gas are the leading sources of energy the world relies on today; and pipelines are viewed as one of the most cost efficient ways to move that energy and deliver it to consumers. The latest data, in 2015, gives a total of more than 3.5 million km of pipeline in 124 countries of the world. Many other thousands of kilometers of pipelines are planned and under construction. Pump stations, along the pipeline, move oil and gas through the pipelines. Because the pipeline walls are under constant pressure, tiny cracks may arise in the steel. Under the continuous load, they can then grow into critical cracks or even leaks. Pipelines conveying flammable or explosive material, such as natural gas or oil, pose special safety concerns; and various accidents have been reported [1]. Damage to the pipeline may cause the occurrence of large and enormous human and economic losses. Moreover, damaged pipelines obviously represent an environmental hazard. Therefore, pipeline operators must identify and remove pipeline failures caused by corrosion and other types of defects as early as possible.

Today, inspection tools, called “Pipeline Inspection Gauges” or “Smart Pigs”, employ complex measuring techniques such as ultrasound and magnetic flux leakage. They are used for the inspection of such pipelines, and have become major components to pipeline safety and accident prevention. These smart pigs are equipped with hundreds of highly tuned sensors that produce data that can be used to locate and determine the thickness of cracks, fissures, erosion and other problems that may affect the integrity of the pipeline. In each inspection passage, huge amounts of data (several hundred gigabytes) are collected. A team of experts will look at these data and assess the health of the pipeline segments.

Because of the size and complexity of pipeline systems and the huge amounts of data collected, human inspection alone is neither feasible nor reliable. Automating the inspection process and the evaluation and interpretation of the collected data have been an important goal for the pipeline industry for a number of years. Significant progress has been made in that regard, and we currently have a number of techniques available that can make the highly challenging and computationally-intensive task of automating pipeline inspection possible. These techniques range from analytical modeling, to numerical computations, to methods employing artificial intelligence techniques such as artificial neural networks. This paper presents a survey of the state-of-the-art in methods used to assess the safety of the oil and gas pipelines, with emphasis on intelligent techniques. The paper explains the principles behind each method, highlights the settings where each method is most effective, and shows how several methods can be combined to achieve higher accuracy.

The rest of the paper is organized as follows. In Sect. 2, we review the five stages of the pipeline reliability assessment process. The theoretical principals behind the intelligent techniques surveyed in this study are discussed in Sect. 3. In

Sect. 4, the pipeline safety assessment approaches using the intelligent techniques reported in Sect. 3 are presented and analyzed. We conclude with final remarks in Sect. 5.

2 Safety Assessment in Oil and Gas Pipelines

The pipeline reliability assessment process is basically composed of five stages, namely data processing, defect detection, determination of defect size, assessment of defect severity, and repair management. Once a defect is detected, the defect assessment unit proceeds by determining the size (the defect's depth and length) of the defect. This is really an important step as the severity of the defect is based on its physical characteristics. Based on the severity level of the detected defect, an appropriate action is taken by the repair management. These five stages of the pipeline assessment process are summarized in the following subsections.

2.1 Big Data Processing

The most common nondestructive evaluation (NDE) method of scanning oil and gas pipelines for possible pipeline defects utilizes magnetic flux leakage (MFL) technology [2], in which autonomous devices containing magnetic sensors are sent on periodic basis into the pipeline under inspection. The magnetic sensors are used to measure MFL signals every three-millimeters along the pipeline length. Figure 1 shows a rolled-out representation of a pipeline wall. The MFL sensors are equally distributed around the circumference of the pipeline and move parallel to the axis of the pipeline.

For pipelines that extend hundreds of kilometers, the data sets collected by the MFL sensors are so big and complex that traditional data processing techniques to analyze such data are inadequate. To reduce the quantity of the data, redundant and

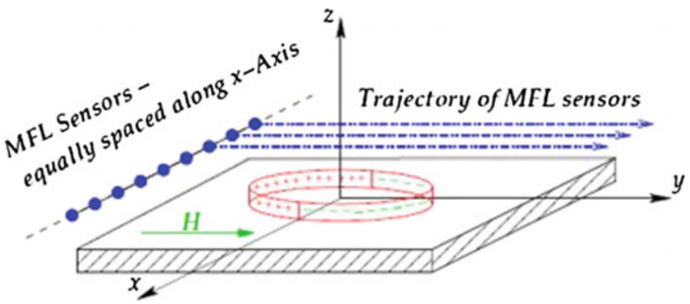

Fig. 1 Rolled-out representation of pipeline wall

irrelevant data are removed using feature extraction and selection techniques. The most relevant features are selected, and then used to determine the depth and length of the detected defect.

2.2 Defect Detection

In this stage, the diagnostic data are examined for the existence of possible defects in the pipeline. To detect and identify the location of potential defects, wavelet techniques are widely used [3]. They are very powerful mathematical methods [4–6]. They were reported in many applications such as data compression [7], data analysis and classification [8], and de-noising [9–11].

2.3 Determination of Defect Size

To determine the severity level of the detected defect, the defect's depth and length are calculated. However, the relationship between the given MFL signals and particular defect type and shape is not well-known. Hence, it is very difficult to derive an analytical model to describe this relationship. To deal with this problem, researchers resort to intelligent techniques to estimate the required parameters. One of these intelligent tools is the Adaptive Neuro-Fuzzy Inference System (ANFIS).

2.4 Assessment of Defect Severity

Based on the defect parameters (i.e., depth and length) obtained in the previous stage, an industry standard known as ASME B31G is often used to assess the severity level of the defect [12]. It specifies the pipeline stress under operating pressure and what defect parameters that may fail the hydro pressure test [13].

2.5 Repair Management

In order to determine an appropriate maintenance action, the repair management classifies the severity level of pipeline defects into three basic categories, namely: severe, moderate, and acceptable. Severe defects are given the highest priority and an immediate action is often required. The other two severity levels are not deemed critical, thus, a repair action can be scheduled for moderate and acceptable defects.

3 Computational Intelligence

As mentioned in the previous section, MFL signals are widely used to determine the depth and length of potential defects. From recorded data, it has been observed that the magnitude of MFL signals varies from one defect depth and length to another. In the absence of analytical models that can describe the relationship between the amplitude of MFL signals and their corresponding defect dimensions, computational intelligence provides an alternative approach. Given sufficient MFL data, there are different computational techniques such as data mining techniques, artificial neural networks, and hybrid neuro-fuzzy systems that can be utilized to learn such relationships. In the following, the theoretical principals behind each of these techniques are summarized.

3.1 Data Mining

The k-nearest neighbor (k-NN) and support vector machines (SVM) are widely used in data mining to solve classification problems. Within the context of the safety assessment in oil and gas pipelines, these two techniques can be employed to assign detected defects to a certain severity level.

3.1.1 K-Nearest Neighbor (KNN)

The KKN is a non-parametric learning algorithm as it does not make any assumptions on the underlying data distribution. This may come in handy since many real world problems do not follow such assumptions. The KNN learning algorithm is also referred to as a lazy algorithm because it does not use the training data points to do any generalization. Thus, there is no training stage in the learning process, but rather KNN makes its decision based on the entire training data set. The learning algorithm assumes that all instances correspond to points in the n-dimensional space. The nearest neighbors of an instance are identified using the standard Euclidean distance. Let us assume that a given defect x is characterized by a feature vector:

$$\langle a_1(x), a_2(x), \ldots, a_n(x) \rangle, \tag{1}$$

where $a_r(x)$ denotes the value of the rth attribute of instance x. Thus, the distance d between two instances x_i and x_j is calculated as follows:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^{n} \left(a_r(x_i) - a_r(x_j)\right)^2}, \tag{2}$$

For the safety assessment in oil and gas pipeline application, the target function is discrete. That is, it assigns the feature vector of the detected defect to one of the three severity levels severe, moderate, or acceptable. If we suppose $k = 1$, then the 1-nearest neighbor assigns the feature vector to the severity level where the training instance of that severity level is nearest to the feature vector. For larger values of k, the algorithm assigns the most common severity level among the k nearest training examples. e only assumption made is that the data is in a feature space.

3.1.2 Support Vector Machine (SVM)

The SVM is a discriminant classifier defined by a separating hyperplane. Given labeled training data, the SVM algorithm outputs an optimal hyperplane that can categorize new examples. Support vector machines are originally designed for binary classification problems. For a linearly separable set of 2D-points, there will be multiple straight lines that may offer a solution to the problem. However, a line is considered bad if it passes too close to the points because it will be susceptible to noise. The task of the SVM algorithm is to find the hyperplane that gives the largest minimum distance (i.e., margin) to the training examples.

To solve multi-class classification problems, the SVM should be extended. The training algorithms of SVMs look for the optimal separating hyperplane which has a maximized margin between the hyperplane and the data, which in turn, minimizes the classification error. The separating hyperplane is represented by a small number of training data, called support vectors (SVs). However, the real data cannot be separated linearly, thus the data are mapped into a higher dimensional space. Practically, a kernel function is utilized to calculate the inner product of the transformed data. The efficiency of the SVM depends mainly on the kernel.

Formally, the hyperplane is defined as follows:

$$f(x) = \beta_0 + \beta^T x, \tag{3}$$

where β is known as the *weight vector* and β_0 as the bias. The optimal hyperplane can be represented in an infinite number of different ways by scaling of β and β_0. The hyperplane chosen is:

$$\left|\beta_0 + \beta^T x\right| = 1, \tag{4}$$

where x symbolizes the training examples closest to the hyperplane, which are called support vectors. The distance between a point x and a hyperplane (β,β_0) can be calculated as:

$$distance = \frac{\left|\beta_0 + \beta^T x\right|}{||\beta||}, \tag{5}$$

For the canonical hyperplane, the numerator is equal to one, thus,

$$distance = \frac{\left|\beta_0 + \beta^T x\right|}{||\beta||} = \frac{1}{||\beta||}, \tag{6}$$

The margin (M) is twice the distance to the closest examples:

$$M = \frac{2}{||\beta||}, \tag{7}$$

Now, maximizing M is equivalent to the problem of minimizing a function $L(\beta)$ subject to some constrains as follows:

$$\min_{\beta, \beta_0} L(\beta) = \frac{1}{2}|\beta|^2, \tag{8}$$

subject to:

$$y_i = \left(\beta^T x_i + \beta_0\right) \geq 1 \forall_i, \tag{9}$$

where y_i represents each of the labels of the training examples.

3.2 Artificial Neural Networks

Artificial neural networks (ANN) are suitable for the safety assessment in oil and gas pipelines as they are capable of solving ill-defined problems. Essentially they attempt to simulate the neural structure of the human brain and its functionality.

The multi-layer perceptron (MLP) with the back propagation learning algorithm is considered the most common neural network and being widely used in a large number of applications. A typical MLP neural network of one hidden layer is depicted in Fig. 2. There are d inputs (example, d dimensions of input pattern X), h hidden nodes, and c outputs nodes.

The output of the jth hidden node is $z_j = f_j(a_j)$, where $a_j = \sum_{i=0}^{d} w_{ji} x_i$, and $f_j(.)$ is an activation function associated with hidden node j. w_{ji} is the connection weight from the input node i to j, and w_{j0} denotes the bias for the hidden node j. For an input node k, its output is $y_k = f_k(a_k)$, where $a_k = \sum_{j=0}^{h} w_{kj} z_j$, and $f_k(.)$ is the activation function associated with output node k. w_{kj} is the connection weight from hidden node j to output node k. w_{k0} denotes the bias for output node k. The activation function is often chosen as the unipolar sigmoidal function:

$$f(a) = \frac{1}{1 + \exp(-\gamma a)}, \tag{10}$$

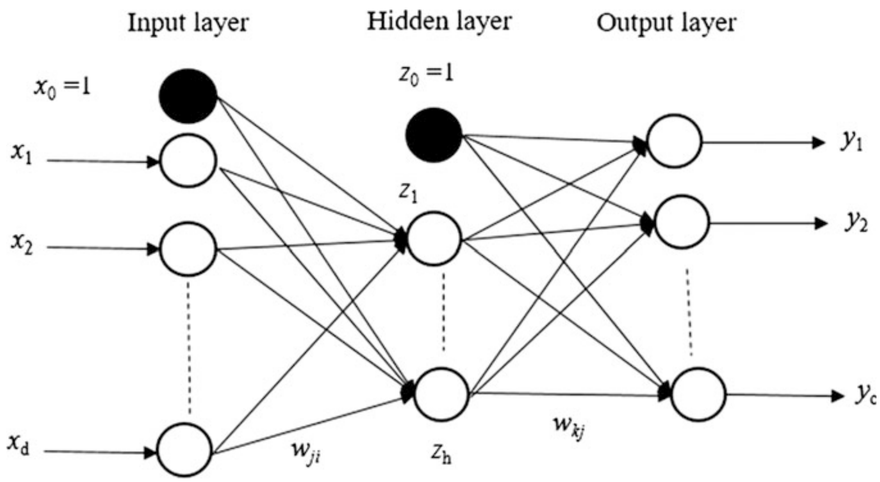

Fig. 2 A multi-layer perceptron neural network

In MLP, the back propagation learning algorithm is used to update weights so as to minimize the following squared error function:

$$J(w) = \frac{1}{2}\sum_{k=1}^{c} (e_k - y_k(X))^2, \tag{11}$$

3.3 *Hybrid Neuro-Fuzzy Systems*

The focus of intelligent hybrid systems in this study will be on the combination of neural networks and fuzzy inference systems. One of these systems is the adaptive neuro-fuzzy inference system (ANFIS), which will be used as an illustrative example of such hybrid systems. ANFIS, as introduced by Jang [14], utilizes fuzzy IF-THEN rules, where the membership function parameters can be learned from training data, instead of being obtained from an expert [15–23]. Whether the domain knowledge is available or not, the adaptive property of some of its nodes allows the network to generate the fuzzy rules that approximate a desired set of input-output pairs. In the following, we briefly introduce the ANFIS architecture as proposed in [14]. The structure of the ANFIS model is basically a feedforward multi-layer network. The nodes in each layer are characterized by their specific function, and their outputs serve as inputs to the succeeding nodes. Only the parameters of the adaptive nodes (i.e., square nodes in Fig. 3) are adjustable during the training session. Parameters of the other nodes (i.e., circle nodes in Fig. 3) are fixed.

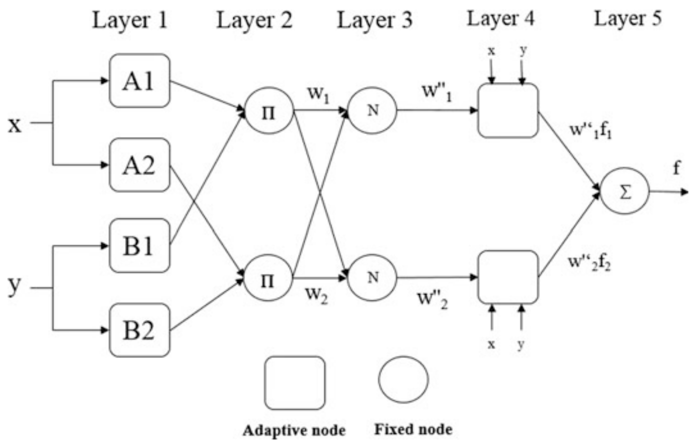

Fig. 3 The architecture of ANFIS

Suppose there are two inputs x, y, and one output f. Let us also assume that the fuzzy rule in the fuzzy inference system is depicted by one degree of Sugeno's function [14].

Rule 1: if x is A_1 and y is B_1 then $f = p_1x + q_1y + r_1$
Rule 2: if x is A_2 and y is B_2 then $f = p_2x + q_2y + r_2$

where p_i, q_i, r_i are adaptable parameters.

The node functions in each layer are described in the sequel.

Layer 1: Each node in this layer is an adaptive node and is given as follows:

$$o_i^1 = \mu_{Ai}(x), \quad i = 1, 2$$
$$o_i^1 = \mu_{Bi-2}(y), \quad i = 3, 4$$

where x and y are inputs to the layer nodes, and A_i and B_{i-2} are linguistic variables. The maximum and minimum of the bell-shaped membership function are 1 and 0, respectively. The membership function has the following form:

$$\mu_{Ai}(x) = \frac{1}{1 + \left\{ \left(\frac{x - c_i}{a_i} \right)^2 \right\}^{bi}}, \quad (12)$$

where the set $\{a_i, b_i, c_i\}$ represents the premise parameters of the membership function. The bell-shaped function changes according to the change of values in these parameters.

Layer 2: Each node in this layer is a fixed node. Its output is the product of the two input signals as follows:

$$o_i^2 = w_i = \mu_{Ai}(x)\mu_{Bi}(y), \quad i = 1, 2, \tag{13}$$

where w_i refers to the firing strength of a rule.

Layer 3: Each node in this layer is a fixed node. Its function is to normalize the firing strength as follows:

$$o_i^3 = w_i^{''} = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \tag{14}$$

Layer 4: Each node in this layer is adaptive and adjusted as follows:

$$o_i^4 = w_i^{''} f_i = w_i^{''}(p_i x + q_i y + r_i), \quad i = 1, 2 \tag{15}$$

where $w_i^{''}$ is the output of layer 3 and $\{p_i + q_i + r_i\}$ is the consequent parameter set.

Layer 5: Each node in this layer is fixed and computes its output as follows:

$$o_i^5 = \sum_{i=1}^{2} w_i^{''} f_i = \frac{\left(\sum_{i=1}^{2} w_i f_i\right)}{w_1 + w_2}, \tag{16}$$

The output of layer 5 sums the outputs of nodes in layer 4 to be the output of the whole network. If the parameters of the premise part are fixed, the output of the whole network will be the linear combination of the consequent parameters, i.e.,

$$f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2, \tag{17}$$

The adopted training technique is hybrid, in which, the network node outputs go forward till layer 4, and the resulting parameters are identified by the least square method. The error signal, however, goes backward till layer 1, and the premise parameters are updated according to the descent gradient method. It has been shown in the literature that the hybrid-learning technique can obtain the optimal premise and consequent parameters in the learning process [14].

4 Pipeline Safety Assessment Using Intelligent Techniques

In this section, pipeline safety assessment approaches using the above intelligent techniques that are reported in the literature are presented and analyzed. Most of these have been proposed for either predicting pipeline defect dimensions or detecting and classifying defect types [24].

4.1 Data Mining-Based Techniques

A recognition and classification of pipe cracks using images analysis and a neuro-fuzzy algorithm is proposed [25]. In the preprocessing step the scanned images of the pipe are analyzed and crack features are extracted. In the classification step the neuro-fuzzy algorithm is developed that employs a fuzzy membership function and an error back-propagation algorithm. The classification of underground pipe defects is carried out using the Euclidean distance method, a fuzzy-KNN algorithm, a conventional back-propagation neural network, and a neuro-fuzzy algorithm. The theoretical backgrounds of all classifiers are presented and their relative advantages are discussed. In conventional recognition methods, the Euclidean distance has been commonly used as a distance measure between two vectors. The Euclidean distance is defined by Eq. 2.

The fuzzy k-NN algorithm assigns class membership to a sample observation based on the observation distance from its k-nearest neighbors and their membership. The neural network universal approximation property guarantees that any sufficiently smooth function can be approximated using a two-layer network. Neuro-fuzzy systems belong to hybrid intelligent systems. Neural networks are good for numerical knowledge (data sets), fuzzy logic systems are good for linguistic information (fuzzy sets). The proposed neuro-fuzzy algorithm is a mixture, where the input and the output of the ANN is a fuzzy entity. Fuzzy neural networks such as the ones proposed in this study provide more flexibility in representing the input space by integrating vagueness usually associated with fuzzy patterns with learning capabilities of neural networks. In fact, by using fuzzy variables as input to the neural network structure, the boundaries of the decision space become represented in a less restrictive manner (unlike the conventional structure of neural networks where the input are required to be crisp), and permits the representation of data possibly belonging to overlapping boundaries. As such more information could be represented without having recourse to the storage of a huge amount of data, which are usually required for the training and testing of conventional "crisp-based data training" neural networks.

The main disadvantage of the KNN algorithm, in addition to determining the value of the parameter k, is that, for a large number of images or MFL data, the computation cost is high because we need to compute the distance of each instance to all training samples. Moreover, it takes up a lot of memory to store all the image properties and features of MFL samples. However, it is simple and effective due to the large data.

SVM-based approaches are reported in [26–28]. In [26], the proposed approach aims at detecting, identifying, and verifying construction features while inspection the condition of underground pipelines. The SVM is used to classify features extracted from the signals of a NDE sensor. The SVM model to be trained for this work uses the RFT data and the ground truth labels to learn how to separate construction features (CF) from other data (non-CF) from CCTV images. The CFs represent pipeline features such as joints, flanges, and elbows. The learned SVM model is later employed to detect CF in unseen data. In [27], the authors propose an SVM method to reconstruct defects shape features. To create a defect feature picture, a large number of samples are collected for each defect. The SVM model reconstruction error is below 4%. For the analysis of magnetic flux leakage images in pipeline inspection, the authors in [28] apply support vector regression among other techniques. In this paper, the focus is on the binary detection problem of classifying anomalous image segments into one of two classes: the first class is the one which consists of injurious or non-benign defects such as various crack-like anomalies and metal losses in girth welds, long-seam welds, or in the pipe wall itself, which if left untreated, could lead to pipeline rupture. The second class consists of non-injurious or benign objects such as noise events, safe and non-harmful pipeline deformations, manufacturing irregularities, etc.

Although finding the right kernel for the SVM classifier is a challenge, but once obtained, it can work well despite the fact that the MFL data is not linearly separable. The main disadvantage is that it is fundamentally a binary classifier; thus, there is no particular way for dealing with multi-defect pipeline problems.

4.2 Neural Network-Based Techniques

Artificial neural networks have been used extensively in safety assessment in oil and gas pipelines [29–33]. In [29], Carvalho et al. propose an artificial neural network approach for detection and classification of pipe weld defects. These defects were manufactured and deliberately implanted. The ANN was able to distinguish between defect and non-defect signals with great accuracy (94.2%). For a particular type of defect signals, the ANN recognized them 92.5% of the time. In [29], a Radial Basis Function Neural Network (RBFNN) is deemed to be a suitable technique and a corrosion inspection tool to recognize and quantify the corrosion characteristics. An Immune RBFNN (IRBFNN) algorithm is proposed to process the MFL data to determine the location and size of the corrosion spots on the pipeline. El Abbasy et al. in [31] propose an artificial neural network models to evaluate and predict the condition of offshore oil and gas pipelines. The inspection data for selected factors are used to train the ANN in order to obtain ANN-based condition prediction models. The inspection data points were divided randomly into three sets: (1) 60% for training; (2) 20% for testing; and (3) 20% for validation. The training set is used to train the network whereas the testing set is used to test the network during the development/training and also to continuously correct it by adjusting the weights of network links. The authors in [32] propose a machine learning approach for big data in oil and gas pipelines, in which three different network architectures are examined, namely static feedforward neural networks (static FFNN), cascaded FFNN, and dynamic FFNN as shown in Figs. 4, 5, and 6, respectively.

In the static FFNN architecture, the extracted feature vector is fed into the first hidden layer. Weight connections, based on the number of neurons in each layer,

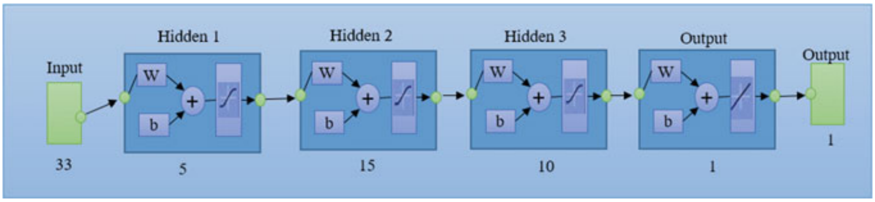

Fig. 4 Architecture of static FFNN

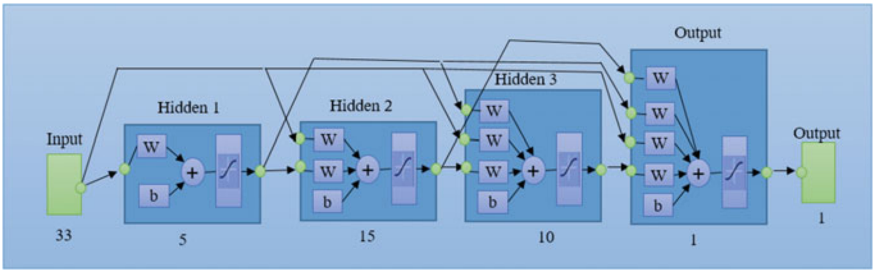

Fig. 5 Architecture of cascaded FFNN

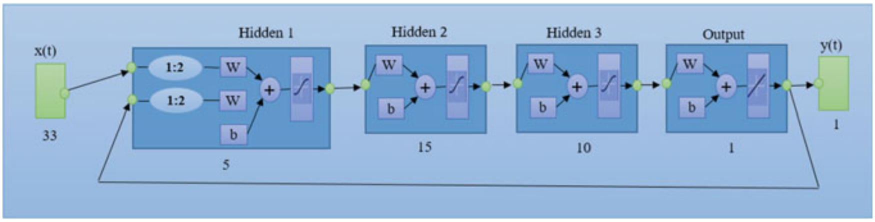

Fig. 6 Architecture of dynamic FFNN

are assigned between every adjacent layers. While in the cascaded FFNN architecture, include a weight connection from the input layer to each other layer, and from each layer to the successive layers. In the dynamic architecture, the network outputs depend not only on the current input feature vector, but also on the previous inputs and outputs of the network. Compared with the performance of pipeline inspection techniques reported by service providers such as GE and ROSEN, the results obtained using the method we proposed are promising. For instance, within $\pm 10\%$ error-tolerance range, the obtained estimation accuracy is 86%, compared to only 80% reported by GE; and within $\pm 15\%$ error-tolerance range, the achieved estimation accuracy is 89% compared to 80% reported by ROSEN.

Mohamed et al. propose a self-organizing map-based feature visualization and selection for defect depth estimation in oil and gas pipelines in [33]. The authors use the self-organizing maps (SOMs) as feature visualization tool for the purpose of

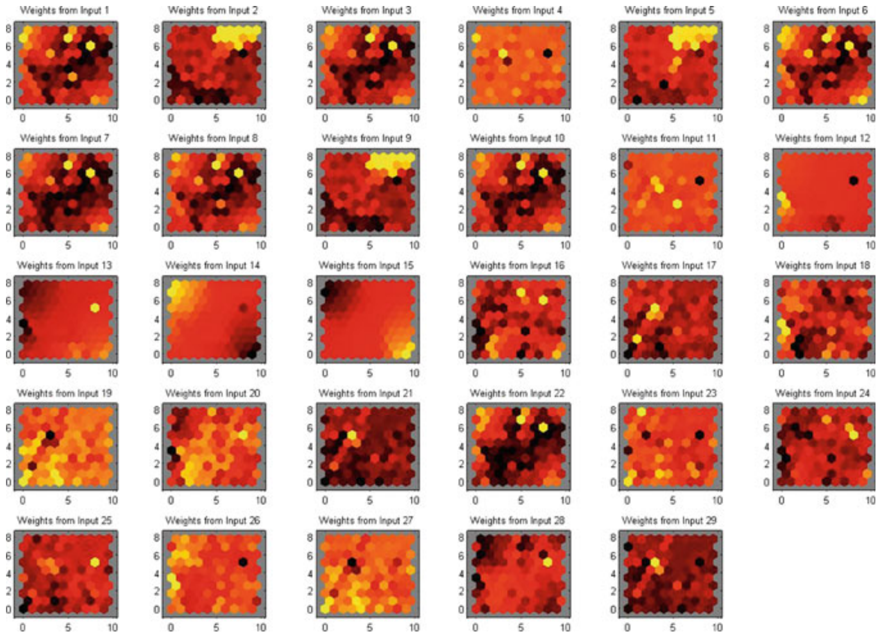

Fig. 7 SOM weights for each input feature [33]

selecting the most appropriate features. The SOM weights for each individual input feature (weight plane) are displayed then visually analyzed. Irrelevant and redundant features can be efficiently spotted and removed. The remaining "good" features (i.e., selected features) are then used as an input to a feedforward neural network for defect depth estimation. An example of the SOM weights are shown in Fig. 7. The 21 features selected by the SOM approach are used to evaluate the performance of the three FFNN structures. Experimental work has shown the effectiveness of the proposed approach. For instance, within ±5% error-tolerance range, the obtained estimation accuracy, using the SOM-based feature selection, is 93.1%, compared to 74% when all input features are used (i.e., no feature selection is performed); and within ±10% error-tolerance range, the obtained estimation accuracy, using the SOM-based feature selection, is 97.5%, compared to 86% when all the input features are used (i.e., no feature selection is performed).

The disadvantage of using neural networks is that the neural network structure (i.e., number of neurons, hidden layers, etc.) is determined by trial and error approach. Moreover, the learning process can take very long due to the large number of MFL samples. The main advantage is that there is no need to find a mathematical model that describes the relationship between MFL signals and pipeline defects.

4.3 Hybrid Neuro-Fuzzy Systems-Based Techniques

Several approaches that utilize hybrid systems have been reported in the literature. In [34], the authors propose a neuro-fuzzy classifier for the classification of defects by extracting features in segmented buried pipe images. It combines a fuzzy membership function with a projection neural network where the former handles feature variations and the latter leads to good learning efficiency as illustrated in Fig. 8. Sometimes the variation of feature values is large, in which case it is difficult to classify objects correctly based on these feature values. Thus, as shown in the figure, the input feature is converted into fuzzified data which are input to the projection neural network. The projection network combines the utility of both the restricted coulomb energy (RCE) network and backpropagation approaches. A hypersphere classifier such as RCE places hyper-spherical prototypes around training data points and adjusts their radii. The neural network inputs are projected onto a hypersphere in one higher dimension and the input and weight vectors are confined to lie on this hypersphere. By projecting the input vector onto a hypersphere in one higher dimension, prototype nodes can be created with closed or open classification surfaces all within the framework of a backpropagation trained feedforward neural network. In general, a neural network passes through two phases: training and testing. During the training phase, supervised learning is used to assign the output membership values ranging in [0,1] to the training input vectors. Each error in membership assignment is fed back and the connection weights of the network are appropriately updated. The back-propagated error is computed with respect to each desired output, which is a membership value denoting the degree of belongingness of the input vector to a certain class. The testing phase in a fuzzy network is equivalent to the conventional network.

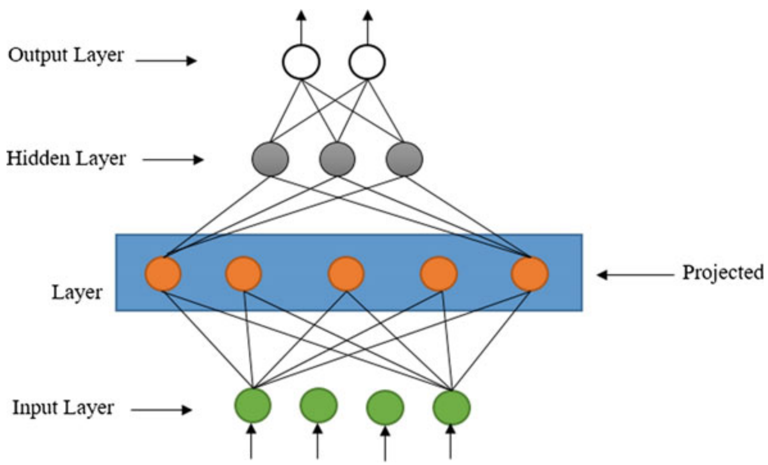

Fig. 8 A hybrid neuro-fuzzy classifier

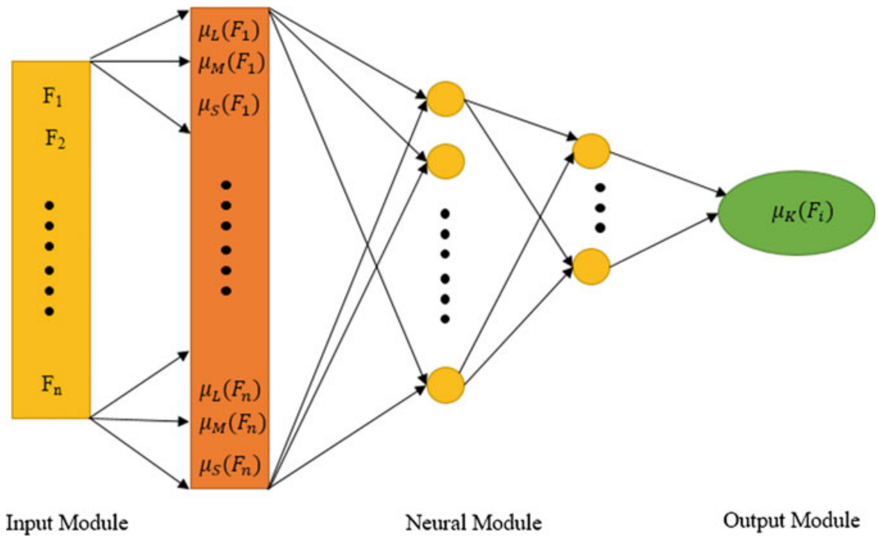

Fig. 9 Neuro-fuzzy neural network architecture

In [35], a classification of underground pipe scanned images using feature extraction and neuro-fuzzy algorithm is proposed. The concept of the proposed fuzzy input and output module and neural network module is illustrated in Fig. 9. The fuzzy ANN model has three modules: the fuzzy input module, the neural network module, and the fuzzy output module. The neural network module is aconventional feedforward artificial neural network. A simple three-layer network with a backpropagation training algorithm is used in this study. To increase the rate of convergence, a momentum term and a modified backpropagation training rule called the delta–delta rule are used. The input layer of this network consists of 36 nodes (because of the use of fuzzy sets to screen the 12 input variables; and the output layer consists of seven nodes (trained with fuzzy output values). As shown in Fig. 9, the input layer of this fuzzy ANN model is actually an output of the input module. On the other hand, the output layer becomes an input to the output module. The input and output modules, for preprocessing and post-processing purposes, respectively, are designed to deal with the data of the ANN using fuzzy sets theory.

In [36], an adaptive neuro-fuzzy inference system (ANFIS)-based approach is proposed to estimate defect depths from MFL signals. To reduce data dimensionality, discriminant features are first extracted from the raw MFL signals. Representative features that characterize the original MFL signals can lead to a better performance for the ANFIS model and reduce the training session. The following features are extracted: maximum magnitude, peak-to-peak distance, integral of the normalized signal, mean average, and standard deviation. Moreover, MFL signals can be approximated by polynomial series of the form, $a_nX^n + \ldots + a_1X + a_0$. The proposed approach is tested for different levels of error-tolerance. At the levels of

±15, ±20, ±25, ±30, ±35, and ±40%, the best defect depth estimates obtained by the new approach are 80.39, 87.75, 91.18, 95.59, 97.06, and 98.04%, respectively.

The advantages of using ANFIS is that the MFL data can be exploited to learn the fuzzy rules required to model the pipeline defects, and it converges faster than typical feedforward neural networks. However, the number of rules extracted is exponential with the number of used MFL features, which may prolong the learning process.

5 Conclusion

In this paper, the applicability of computational intelligence in the safety assessment in oil and gas pipelines is surveyed and examined. The survey covers safety assessment approaches that utilize data mining techniques, artificial neural networks, and hybrid neuro-fuzzy systems, for the purpose of detecting pipeline defects, estimating their dimensions, and identifying (classifying) their severity level. Obviously, techniques of computational intelligence offer an attractive alternative to traditional approaches as they can cope with complexity resulting from the uncertainty accompanying the collected diagnostic data, as well from the large size of the collected data. For intelligent techniques such as KNN, SVM, neural networks, and ANFIS, there is no need to derive a mathematical model that describes the relationship between pipeline defects and the diagnostic data (i.e., MFL and ultra sound signals, images, etc.). For typically large MFL data, KNN and SVM classifiers perform well and can provide optimal results. However, KNN may require large memory to store MFL samples. Obtaining suitable kernel functions for the SVM model has proven to be difficult. While, large MFL data may effectively be used to train different types and structures of neural networks, the learning process may take long time. Moreover, appropriate fuzzy rules can be extracted from the MFL data for the ANFIS model, which has the advantage of converging much faster than regular neural networks. The number of rules extracted, however, may increase exponentially with the number of the used MFL features.

References

1. http://www.ntsb.gov/investigations/AccidentReports/Pages/pipeline.aspx
2. K. Mandal and D. L. Atherton, A study of magnetic flux-leakage signals. Journal of Physics D: Applied Physics, 31(22): 3211, 1998.
3. K. Hwang, et al, Characterization of gas pipeline inspection signals using wavelet basis function neural networks. NDT & E International, 33(8): 531-545, 2000.
4. I. Daubechies, Ten Lectures on Wavelets. SIAM: Society for Industrial and Applied Mathematics, 1992.
5. S. Mallat, A Wavelet Tour of Signal Processing. Academic Press, 2008.

6. M. Misiti, Y. Misiti, G. Oppenheim, and J. M. Poggi, Wavelets and their Applications. Wiley-ISTE, 2007.
7. J. N. Bradley, C. Bradley, C. M. Brislawn, and T. Hopper, FBI wavelet/scalar quantization standard for gray-scale fingerprint image compression. In: SPIE Procs, Visual Information Processing II, 1961: 293–304, 1993.
8. M. Unser and A. Aldroubi, A review of wavelets in biomedical applications. Proceedings of the IEEE, 84 (4): 626–638, 1996.
9. S. Shou-peng and Q. Pei-wen, Wavelet based noise suppression technique and its application to ultrasonic flaw detection. Ultrasonics, 44(2): 188–193, 2006.
10. A. Muhammad and S. Udpa, Advanced signal processing of magnetic flux leakage data obtained from seamless gas pipeline. Ndt & E International, 35(7): 449–457, 2002.
11. S. Mukhopadhyay and G. P. Srivastava, Characterization of metal loss defects from magnetic flux leakage signals with discrete wavelet transform. NDT & E International, 33(1): 57–65, 2000.
12. American Society of Mechanical Engineers, ASME B31G Manual for Determining Remaining Strength of Corroded Pipelines, 1991.
13. A. Cosham and M. Kirkwood, Best practice in pipeline defect assessment. Paper IPC00–0205, International Pipeline Conference, 2000.
14. J. R. Jang, ANFIS: Adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, MAN, and Cybernetics, 23: 665–685, 1993.
15. N. Roohollah, S. Safavi, and S. Shahrokni, A reduced-order adaptive neuro-fuzzy inference system model as a software sensor for rapid estimation of five-day biochemical oxygen demand. Journal of Hydrology, 495: 175–185, 2013.
16. K. Mucsi, K. Ata, and A. Mojtaba, An Adaptive Neuro-Fuzzy Inference System for estimating the number of vehicles for queue management at signalized intersections. Transportation Research Part C: Emerging Technologies, 19(6): 1033–1047, 2011.
17. A. Zadeh, et al, An emotional learning-neuro-fuzzy inference approach for optimum training and forecasting of gas consumption estimation models with cognitive data. Technological Forecasting and Social Change, 91: 47–63, 2015.
18. H. Azamathulla, A. Ab Ghani, and S. Yen Fei, ANFIS-based approach for predicting sediment transport in clean sewer. Applied Soft Computing, 12(3): 1227–1230, 2012.
19. A. Khodayari, et al, ANFIS based modeling and prediction car following behavior in real traffic flow based on instantaneous reaction delay. IEEE Intelligent Transportation Systems Conference, 599–604, 2010.
20. A. Kulaksiz, ANFIS-based estimation of PV module equivalent parameters: application to a stand-alone PV system with MPPT controller. Turkish Journal of Electrical Engineering and Computer Science, 21(2): 2127–2140, 2013.
21. D. Petković, et al, Adaptive neuro-fuzzy estimation of conductive silicone rubber mechanical properties. Expert Systems with Applications, 39(10): 9477–9482, 2012.
22. M. Chen, A hybrid ANFIS model for business failure prediction utilizing particle swarm optimization and subtractive clustering. Information Sciences, 220: 180–195, 2013.
23. M. Iphar, ANN and ANFIS performance prediction models for hydraulic impact hammers. Tunneling and Underground Space Technology, 27(1): 23–29, 2012.
24. Layouni, Mohamed, Sofiene Tahar, and Mohamed Salah Hamdi, "A survey on the application of neural networks in the safety assessment oil and gas pipelines." Computational Intelligencefor Engineering Solutions", 2014 IEEE Symposium on. IEEE, 2014.G. S. Park, and E. S. Park, Improvement of the sensor system in magnetic flux leakage-type nod-destructive testing. IEEE Transactions on Magnetics, 38(2): 1277–1280, 2002.
25. Sinha, Sunil K., and Fakhri Karray. "Classification of underground pipe scanned images using feature extraction and neuro-fuzzy algorithm." IEEE Transactions on Neural Networks 13.2 (2002): 393–401.R. K. Amineh, et al, A space mapping methodology for defect characterization from magnetic flux leakage measurement. IEEE Transactions on Magnetics, 44(8): 2058-2065, 2008.

26. Vidal-Calleja, Teresa, et al. "Automatic detection and verification of pipeline construction features with multi-modal data." 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2014.
27. Lijian, Yang, et al. "Oil-gas pipeline magnetic flux leakage testing defect reconstruction based on support vector machine." Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on. Vol. 2. IEEE, 2009.
28. Khodayari-Rostamabad, Ahmad, et al. "Machine learning techniques for the analysis of magnetic flux leakage images in pipeline inspection." IEEE Transactions on magnetics 45.8 (2009): 3073–3084.
29. Carvalho, A. A., et al. "MFL signals and artificial neural networks applied to detection and classification of pipe weld defects." Ndt & E International 39.8 (2006): 661–667.
30. Ma, Zhongli, and Hongda Liu. "Pipeline defect detection and sizing based on MFL data using immune RBF neural networks." Evolutionary Computation, 2007. CEC 2007. IEEE Congress on. IEEE, 2007.
31. El-Abbasy, Mohammed S., et al. "Artificial neural network models for predicting condition of offshore oil and gas pipelines." Automation in Construction 45 (2014): 50–65.
32. A. Mohamed, M. S. Hamdi, and S. Tahar, A machine learning approach for big data in oil and gas pipelines. 3rd International Conference on Future Internet of Things and Cloud (FiCloud), IEEE, 2015.
33. A. Mohamed, M. S. Hamdi, and S. Tahar, Self-organizing map-based feature visualization and selection for defect depth estimation in oil and gas pipelines. 19th International Conference on Information Visualization (iV), IEEE, 2015.
34. Sinha, Sunil K., and Paul W. Fieguth. "Neuro-fuzzy network for the classification of buried pipe defects." Automation in Construction 15.1 (2006): 73–83.
35. Sinha, Sunil K., and Fakhri Karray. "Classification of underground pipe scanned images using feature extraction and neuro-fuzzy algorithm." IEEE Transactions on Neural Networks 13.2 (2002): 393–401.
36. A. Mohamed, M. S. Hamdi, and S. Tahar, An adaptive neuro-fuzzy inference system-based approach for oil and gas pipeline defect depth estimation. SAI Intelligent Systems Conference. IEEE, 2015.