Optimized learning by artificial neural network for inspection of heating equipment with infrared sensor data

T Pham

Professor & Director, Centro de Investigación en Tecnología de Información, Universidad de Talca, Talca, Chile

E-mail: tpham@utalca.cl

Abstract. Equipment inspection is normally done through the analysis of sensor data measured at various points in the equipment. The analysis of data will lead to one of two states: pass, or fail. In this scenario, using a trained artificial neural network can bypass the requirement of expert knowledge of how the data relate to the internal functions of a piece of equipment. However, neural network has its limitation of a linearized approximation of often nonlinear phenomenon. Furthermore, the sequence of data feeding into the training process can easily affect the knowledge gained through the training. In this paper, the two weaknesses of linearization and dependence on the sequence that data are fed into the training process are demonstrated. A solution of global optimization is proposed to address these weaknesses. The solution is applied to the inspection of heaters that provide heat with hot liquid that moves through a convoluted pipe while radiating heat, where data is obtained in the form of infrared images.

1. Introduction

Machine learning [1] has been an interesting topic in the discipline of artificial intelligence where learning is implemented in computers. In a setting of supervised learning [2], a training session is conducted so that a general pattern is recognized and recorded so that it can be applied to future similar scenario. This supervised learning is commonly implemented in an artificial neural network [3] where a pattern is coded into a set of constants characterizing the processing of input data to output result similar to that of an actual neuron. The set of constants representing a pattern is obtained when known pairs of input and output are fed into the network. Thus, detailed expertise of a process or system is not necessary, and its behaviour is still represented with proper training.

In this paper, a routine task of inspection of equipment is implemented into a simple perceptron (a special neuron) that takes a set of input data and produces an output of value 1 or −1, where the value 1 represents pass state, and −1 fail state. The perceptron model should function with relative ease when proper training is conducted. However, in the implementation where the classical Adaline learning [4] is used because of its simplicity, the resulting pattern was not obtained correctly. It was discovered through trial-end-error experiments that the resulting pattern is dependent on the order in which the data are fed to the perceptron, i.e., during training the same set of data that is fed into a perceptron in different order will produce different results. Furthermore, when there are overlapping
data due to nonlinear property of a process or system, the results are even more erratic and often unstable.

The inspection of equipment should be easily done with the classical Adaline learning. However, the constants in the perceptron are updated in a localized reactionary manner and therefore do not represent the global optimization \[5\] of the error. In this paper, this learning is modified with the solving of an optimization problem where the square normed \[6\] error is minimized. The result is a unique solution \[7\] that is not dependent on the order in which data are fed sequentially during the training session. Furthermore, even in the nonlinear scenario where there are overlapping data, the result still converges \[8\] correctly. The technique is applied to the inspection of heaters with hot liquids moving through a convoluted pipe while radiating heat to the environment where heat is measured through the use of infrared images and fed into the perceptron to obtain a quick inspection.

2. **Approximating a Process with Linear Model in a Perceptron**

A perceptron is a simplest neuron model that takes a number of input data, combines them in a weighted linear combination, and passes this intermediate result to a threshold filter produce a discrete output that has a value of 1 or \(-1\). Figure 1 illustrates this perceptron, and figures 2 and 3 show the erratic behaviour observed with the same set of training data that is fed sequentially during the training session. Furthermore, even in the nonlinear scenario where there are overlapping data, the result still converges \[8\] correctly. The technique is applied to the inspection of heaters with hot liquids moving through a convoluted pipe while radiating heat to the environment where heat is measured through the use of infrared images and fed into the perceptron to obtain a quick inspection.

Considering the perceptron depicted in figure 1 with two inputs \((x \text{ and } y)\) and one output \((z)\), the training is designed to find a straight line \(w_1x + w_2y + \theta = 0\) that separates two given clusters of data from each other. For simplicity without loss of generality, the constant \(w_2\) is normalized to 1. The special case where \(w_2 = 0\) for a vertical line will be addressed later in this section. Let a set of training data \(S = \{ p_n = (x_n, y_n) | n = 1, 2, ..., N \}\) and the corresponding discrimination data \(d(x_n, y_n)\) for \(n = 1, 2, ..., N\) are available, the training task of figuring out the constants \(w_1\) and \(\theta\) can be set up as solving an optimization problem that minimizes the square normed error function:

\[
\min_{w_1, \theta} \sum_{n=1}^{N} (y_n + w_1x_n + \theta)^2 \left[ \text{sgn}(y_n + w_1x_n + \theta) - d(x_n, y_n) \right]^2,
\]

where the sign function \(\text{sgn}(\cdot)\) returns +1 if its argument is positive, 0 if its argument is zero, and \(-1\) if its argument is negative. Notice that the use of the expression in the square brackets of (1) allows the selective consideration of only the data that caused the error while discarding the data that did not cause any error. Since (1) is convex, there exists a solution and this solution is unique \[10\]. Taking the partial derivatives of (1) with respect to \(w_1\) and \(\theta\), and set them to equal to zero as the first step:
∑_{n=1}^{N} 2x_n (y_n + w_1 x_n + \theta) \left[ \text{sgn}(y_n + w_1 x_n + \theta) - d(x_n, y_n) \right]^2 + \sum_{n=1}^{N} 2(y_n + w_1 x_n + \theta)^2 \left[ \text{sgn}(y_n + w_1 x_n + \theta) - d(x_n, y_n) \right] \delta(y_n + w_1 x_n + \theta) = 0, \quad (2a) \\
\sum_{n=1}^{N} 2(y_n + w_1 x_n + \theta)^2 \left[ \text{sgn}(y_n + w_1 x_n + \theta) - d(x_n, y_n) \right] \delta(y_n + w_1 x_n + \theta) = 0, \quad (2b)

where \( \delta(\cdot) \) is the impulse function that is equal to zero everywhere except at zero where it is approaching infinity. The parts involving the impulse function can be discarded because it is always zero when the error occurs (selective condition in the expression within the square brackets), and consequently (2a) and (2b) can be solved simultaneously to yield the solution:

\[
\begin{bmatrix} w_1 \\ \theta \end{bmatrix} = \left[ \sum_{n=1}^{N} x_n \varphi_n \right]^{-1} \left[ \sum_{n=1}^{N} x_n y_n \varphi_n \right] = \frac{1}{\Delta} \begin{bmatrix} \sum_{n=1}^{N} \varphi_n \\ -\sum_{n=1}^{N} x_n \varphi_n \\ \sum_{n=1}^{N} \varphi_n \\ \sum_{n=1}^{N} y_n \varphi_n \end{bmatrix}, \quad (3a)
\]

\[
\Delta = \sum_{n=1}^{N} x_n^2 \varphi_n \sum_{n=1}^{N} \varphi_n - \left( \sum_{n=1}^{N} x_n \varphi_n \right)^2, \quad (3b)
\]

where \( \varphi_n = \frac{1}{4} \left[ \text{sgn}(y_n + w_1 x_n + \theta) - d(x_n, y_n) \right]^2 \). Notice that in (3a), when the inversion of a singular matrix occurs, i.e., \( \Delta = 0 \) and (3a) does not yield a solution, it means that \( w_2 \) is equal to zero and cannot be normalized. In this special instance, the solution is obtained in similar manner by normalizing \( w_1 \) to one and solving for \( \theta \) (with the value of \( w_2 \) expected to be numerically zero):
\[
\begin{bmatrix}
W_2
\theta
\end{bmatrix}
= \left[ \sum_{n=1}^{N} y_n^2 \phi_n \sum_{n=1}^{N} y_n \phi_n \right]^{-1}
\left[ \sum_{n=1}^{N} x_n y_n \phi_n \right]
= \frac{1}{\Delta N}
\left[ \sum_{n=1}^{N} \phi_n - \sum_{n=1}^{N} y_n \phi_n \right]
\sum_{n=1}^{N} x_n y_n \phi_n.
\]

(3c)

The solution in (3) will be repeated into an iterative manner until there are no further changes in the solution. In this iterative algorithm, notice that the nominal solution \(w_1(0)\) and \(b(0)\) can affect the convergence, and therefore must be initialized as close to the final solution as possible. In this aspect, the nominal solution of a line perpendicular to the segment connecting the centers \((x_{C1}, y_{C1})\) and \((x_{C2}, y_{C2})\) of the two clusters \(C_1\) and \(C_2\) at the center \((\frac{1}{2}(x_{C1}+x_{C2}), \frac{1}{2}(y_{C1}+y_{C2}))\) of this segment is proposed for the simple reason that it is the maximum likelihood solution \([9]\) when the two clusters are normally distributed. In this scheme, the nominal solution is:

\[
w_1^{(0)} = \frac{x_2 - x_1}{y_2 - y_1}, \quad \text{and} \quad \theta^{(0)} = -\frac{1}{2}(y_1 + y_2) - \frac{x_2 - x_1}{2(y_2 - y_1)}(x_1 + x_2).
\]

(4)

3. Application in the Inspection of Heating Equipment

The inspection of the type of heaters that use hot liquid channelled through a convoluted pipe where the heat in the hot liquid is radiated to the air can be done with infrared camera (figure 4). In this infrared image, the temperature of the convoluted pipe is measured at various places. Typically, the temperatures at the entry point where hot liquid is flowing into and at the exit point where cooled liquid is flowing out are measured. If heat dissipates correctly, the temperature at the exit point should be lower than the temperature at the entry point. While the rate of fluid flowing is important, it is not detected with infrared sensor, and will be checked at the meter measuring the flow separately. Here it is desirable to look at the measurement and quickly determine if heat dissipates correctly. Figure 5 shows the results with a few training data points to verify the workability of the learning algorithm. Figure 6 shows results with more data for training. Figure 7 shows the results with an undefined gap between failure and success. Figures 8 shows the results with an overlapping region where experts differ in their opinion. Figure 9 shows the results with a nonlinear boundary.

4. Conclusion

A new learning algorithm is derived for the simple perceptron model that is commonly used in artificial neural network. Regardless of the order that data is used, the solution is always unique and optimizes the square normed of the error. This simple perceptron model is used in the inspection of heating equipment, where the training data always yields unique results that optimizes the error. Various training sets are used to illustrate the workability of this new learning algorithm.

References

[1] Murphy K P 2012 Machine Learning: A Probabilistic Perspective (Cambridge, MA: MIT Press)
[2] Suresh S, Sundararajan N and Savitha R 2013 Supervised Learning with Complex-Valued Neural Networks (New York, NY: Springer)
[3] Hassoun M 2003 Fundamentals of Artificial Neural Networks (Cambridge, MA: MIT Press)
[4] Yeung D S, Cloete I, Shi D and Ng W W Y 2010 Sensitivity Analysis for Neural Networks (New York, NY: Springer)
[5] Rao S S 2009 Engineering Optimization: Theory and Practice (Hoboken, NJ: Wiley)
[6] Trefethen L N 2012 Approximation Theory and Approximation Practice (Philadelphia, PA: SIAM)
[7] Bertsekas D P 2015 Convex Optimization Algorithms (Nashua, NH: Athena Scientific)
[8] Bertsekas D 2016 Nonlinear Programming (Nashua, NH: Athena Scientific)
[9] Krose B and Smagt P 1996 An Introduction to Neural Networks (Amsterdam: The University of Amsterdam)
[10] Lax P D (2002) Functional Analysis (New York, NY: Wiley-Interscience)