ISSN 2007-9737

# Detection of Silent Water Leaks in Household Using Artificial Intelligence Methods

Uriel Amado Ramírez-Hernández, Francisco Rafael Trejo-Macotela*, Daniel Robles-Camarillo, Jorge A. Ruiz-Vanoye, Eric Simancas-Acevedo, Rocío Ortega-Palacios, Julio Cesar Ramos-Fernandez

Polytechnic University of Pachuca, Research and Postgraduate Studies, Mexico

trejo_macotela@upp.edu.mx

**Abstract.** Water losses in distribution systems constitute a significant global challenge, undermining water resource sustainability, increasing operational costs, and threatening the water security of millions. In Latin America, up to 40% of treated water is reportedly lost due to leaks, ruptures, and defective connections (Xylem, 2025). At the household level, silent leaks—particularly in toilet flushing systems—can waste over 37,850 litres annually per dwelling (US EPA, 2024). Various international studies have addressed early leak detection using intelligent systems. In Europe, wireless sensor networks and machine learning models such as Random Forest, Support Vector Machines, and neural networks have been deployed for anomaly detection in urban networks. Asian research has demonstrated detection accuracies exceeding 97% through convolutional neural networks trained on acoustic and vibrational signals, enhanced by contrastive learning to address data scarcity. Hybrid approaches combining hydraulic modelling with AI have been applied in the Middle East and China, whereas logic-based and anomaly detection algorithms have been integrated into real-time platforms in Australia and Canada. Sensor placement optimisation via graph partitioning has further improved coverage efficiency. Despite their effectiveness, these solutions often require substantial investment and advanced infrastructure, limiting their applicability in resource-constrained environments. This study proposes a cost-effective, perceptron-based model for detecting silent leaks in household toilets, integrated within an Internet of Things (IoT) framework. The system employs a Hall-effect flow sensor to capture high-resolution filling-time and pulse-count data, processed through supervised learning to discriminate between normal consumption and leakage. Experimental results under real-use conditions achieved 98% classification accuracy, demonstrating both technical feasibility and operational suitability. This approach offers a practical, computationally efficient solution for domestic contexts in Latin America, enabling real-time monitoring and immediate user alerts, thus supporting water conservation efforts through accessible intelligent detection.

**Keywords:** Water leaks, prediction models, water conservation, IoT, smart water, algorithms.

## 1 Introduction

Water scarcity and the sustainable management of potable water resources have become pressing concerns on a global scale, affecting both industrialised and developing nations. Losses within distribution networks represent one of the most significant challenges to achieving water security, with far-reaching economic, environmental, and social implications. Recent estimates indicate that, in many Latin American cities, up to 40% of treated water is lost through leaks, ruptures, and defective connections (Xylem, 2025). While large-scale network failures attract considerable attention, small and often undetected leaks within domestic systems, such as those in toilet discharge mechanisms, can contribute disproportionately to annual water wastage—sometimes exceeding 37,850 litres per household (US EPA, 2024).

In recent decades, a diverse range of methodologies for leak detection has emerged, leveraging developments in sensor technology, data analytics, and artificial intelligence (AI). European initiatives have adopted wireless sensor networks combined with advanced machine learning algorithms—including Random Forest, Support Vector Machines (SVM), and deep neural architectures—to identify anomalies in pressure and flow profiles across urban infrastructures. Parallel efforts in Asia have yielded high-accuracy solutions through convolutional neural networks (CNNs) trained on acoustic and vibrational signals, benefiting from contrastive learning techniques to mitigate the limitations of scarce labelled data (Liu et al., 2020). Hybrid approaches, integrating hydraulic modelling with AI, have been successfully applied in regions such as the Middle East and China, enhancing leak localisation in complex topologies (Fereidooni et al., 2020).

In countries including Australia and Canada, logic-based strategies have been combined with anomaly detection algorithms to build real-time monitoring platforms, capable of issuing alerts and facilitating rapid intervention (Joseph et al., 2024). Meanwhile, research on optimal sensor deployment using graph partitioning has demonstrated that strategic placement can achieve maximum coverage with minimal infrastructure (Hu et al., 2021). Despite their proven effectiveness, many of these solutions are dependent on substantial investments, sophisticated infrastructure, or specific operational conditions, making them less accessible in low-resource environments.

The need for more accessible, locally adaptable, and computationally efficient solutions has therefore become evident, particularly in the context of domestic water systems in Latin America. Recent regional studies, such as those undertaken in Chile and Peru, have begun to employ artificial neural networks—often trained with both simulated and real datasets—to detect leaks with encouraging results (Constanzo & Ignacio, 2022; Apesteguía & Huarcaya, 2018).

Against this backdrop, the present research focuses on the development and evaluation of a simplified yet effective approach: a perceptron-based model designed for the detection of silent leaks in household toilets, supported by Internet of Things (IoT) flow sensing. This strategy aims to balance detection accuracy with ease of deployment, offering a viable pathway towards widespread adoption in residential contexts where the integration of cost-effective, intelligent monitoring could yield substantial water conservation benefits.

## 2 State of the Art

The detection of leaks within water distribution systems has evolved into a multidisciplinary research domain, drawing on advances in fluid mechanics, signal processing, and computational intelligence. Traditional methods—such as manual inspections, visual verification, and acoustic correlation—remain relevant in certain contexts; however, they are frequently limited by delayed detection, insufficient sensitivity, and high operational costs (Alves Coelho et al., 2020). Over the past two decades, the field has experienced a marked transition towards automated, data-driven approaches, leveraging sensor networks, real-time monitoring, and algorithmic analysis to achieve higher accuracy, faster response, and greater scalability (Fan et al., 2022).

The deployment of advanced sensing technologies capable of capturing diverse data modalities—pressure, flow, vibration, and acoustic signatures—has been a principal driver of innovation. These data streams are processed using a wide spectrum of computational techniques, ranging from classical machine learning algorithms, such as decision trees, Random Forest, and Support Vector Machines, to more sophisticated deep learning architectures, including convolutional neural networks and hybrid models integrating physical hydraulic simulations with AI-based inference (Fereidooni et al., 2020; Guo et al., 2021). Integration with Supervisory Control and Data Acquisition (SCADA) systems has enabled continuous monitoring of extensive networks under varied operational conditions (Joseph et al., 2024).

In contexts where labelled data are scarce, recent studies have demonstrated the effectiveness of contrastive learning and data augmentation in improving robustness and generalisability, even in noisy environments (Liu et al., 2020). Furthermore, optimisation strategies for sensor placement—based on graph partitioning—have been shown to maximise detection coverage while minimising instrumentation requirements (Hu et al., 2021). Despite these advances, many state-of-the-art approaches require substantial investment, sophisticated infrastructure, or specialised conditions, thereby limiting their applicability in domestic or resource-constrained settings. This constraint underscores the necessity for simplified, computationally efficient solutions capable of operating reliably in household environments while maintaining high detection accuracy.

## 2.1 Algorithms Applied to Water Leak Detection

Leaks are undesirable phenomena in water distribution systems. The process for their early detection at precise locations within distribution networks is performed through various advanced algorithmic techniques, primarily based on artificial intelligence, machine learning, and signal processing methodologies. The most effective approaches combine water sensors such as ultrasonic devices, SCADA systems, data analytics, and predictive models to identify and locate leaks with greater accuracy.

## 2.2 Algorithms and Methods Employed

Algorithms such as decision trees, random forests, neural networks, support vector machines (SVM), and deep learning methods (Deep Neural Networks, CNNs) are utilised to analyse sensor data and detect patterns associated with leaks.

In 2020, researchers João Alves Coelho, André Glória, and Pedro Sebastião from ISCTE – University Institute of Lisbon (Portugal) published the article entitled *"Precise Water Leak Detection Using Machine Learning and Real-Time Sensor Data"* in the scientific journal *IoT*. In this study, the authors proposed an innovative system for the accurate detection of water leaks in distribution networks, employing a network of wireless sensors to collect real-time data. These data were analysed through machine learning algorithms, including random forests, decision trees, neural networks, and SVMs, to identify and localise leaks with high precision. The system was validated through a real-world case study, demonstrating a leak detection accuracy of 75%, underscoring its effectiveness and potential for improving water resource management and conservation (Alves Coelho et al., 2020).

In November 2023, Suan Lee and Byeonghak Kim, from the School of Computer Science at Semyung University, South Korea, published the article *"Machine Learning Model for Leak Detection Using Water Pipeline Vibration Sensor"* in the journal *Sensors*. Their study introduced a leak detection system based on the analysis of vibrations in pipelines, captured via vibration sensors installed in water meter boxes and valve rooms. Initially, the signals were transformed into tabulated frequency band data. Subsequently, multiple machine learning models were compared, including KNN, decision trees, random forest, extra trees, LightGBM, XGBoost, and CatBoost. The XGBoost model achieved the highest accuracy, with an impressive 99.79%. Additionally, the study differentiated between the behaviour of metallic (better vibration transmission) and non-metallic pipes, further reinforcing the robustness of the approach (Lee & Kim, 2023).

In 2023, researchers Jungyu Choi and Sungbin Im from Soongsil University, South Korea, developed an automatic leak detection system for water pipelines using convolutional neural networks (CNNs). Their study was published in the journal *Applied Sciences*. The method is based on capturing vibrational sound through sensors installed on pipes. These acoustic signals are transformed into magnitude spectra (e.g., FFT), which serve as input for a specifically designed CNN capable of distinguishing between normal conditions and the presence of leaks. The model achieved an F1-score of 94.82% and a Matthews correlation coefficient (MCC) of 94.47%, significantly outperforming traditional SVM models. This approach proved to be more accurate, robust, and effective in automated leak detection, reducing the need for human intervention and enhancing early incident response in water networks (Choi & Sungbin Im, 2023).

In 2022, Guancheng Guo and his team from Tsinghua University (China) proposed the Time–Frequency Convolutional Neural Network (TFCNN) model for detecting leaks in water distribution networks. The method converts acoustic signals into spectrograms representing leakage patterns and trains a CNN to recognise them even under high noise conditions. The model was evaluated against classical algorithms such as SVM, decision trees, and XGBoost, demonstrating an average accuracy of 98%, and up to 90% accuracy in low signal-to-noise ratio (SNR) environments. In real-world trials, it achieved up to 99% accuracy. In conclusion, the TFCNN approach stands out for its robustness, precision, and adaptability, being highly effective for acoustic leak detection in noisy settings (Guo et al., 2021).

In 2024, researchers from Hangzhou Dianzi University (China) and other institutions developed a continuous water leak detection system based on logarithmic spectrograms and convolutional neural networks (CNNs). Acoustic signals captured by network sensors were transformed into logarithmic spectrograms to highlight leakage patterns in the time-frequency domain. The CNN trained on this data achieved high accuracy and sensitivity, even under noisy conditions. This approach proved effective for real-time monitoring and early leak detection, offering a robust solution for enhancing water infrastructure management (Peng et al., 2024).

In 2023, Harshit Shukla and Kalyan R. Piratla from Clemson University (USA) developed a machine learning-based model for leak detection in water distribution networks. The approach proposed an end-to-end classification system wherein feature extraction and leak detection are performed automatically and simultaneously.

The model was applied to both simulated and real hydraulic networks, achieving a detection accuracy of 95.6%, rendering it more effective than traditional methods. Its implementation allows for faster response to leaks and improved predictive maintenance management, contributing to the operational efficiency of urban water infrastructures (Shukla & Piratla, 2023).

In 2021, T. Ravichandran, K. Gavahi, K. Ponnambalam, and collaborators published an article entitled *"Ensemble-Based Machine Learning Approach for Improved Leak Detection in Water Mains"*. This study focuses on applying ensemble learning techniques to enhance leak detection in water distribution networks. The proposed approach combines multiple predictive models to increase the accuracy of leak identification, surpassing the limitations of traditional methods. By integrating various machine learning algorithms, the system can analyse real-time data from sensors installed on pipelines, enabling faster and more accurate detection of anomalies potentially indicating leaks. This advancement represents a significant contribution to the field of water resource management, offering an effective tool to reduce water losses and improve the operational efficiency of distribution networks (Ravichandran et al., 2021).

Contrastive learning enhances detection performance when labelled data are scarce, outperforming traditional supervised methods in data-limited scenarios.

In November 2024, Rongsheng Liu, Tarek Zayed, and Rui Xiao from The Hong Kong Polytechnic University published the article *"Contrastive Learning Method for Leak Detection in Water Distribution Networks"* in the journal *npj Clean Water*. In this study, acoustic signals from water distribution networks were collected using specialised sensors.

To improve the model's training capacity with unlabelled data, data augmentation techniques were applied, revealing that a combination of flip-x and amplitude scaling yielded the best results within the contrastive learning framework. Subsequently, they trained a CNN comprising five convolutional blocks, whose architecture was validated through ablation experiments and t-SNE analysis to ensure optimal performance.

The contrastive learning-based model demonstrated superior performance compared to supervised approaches, particularly under conditions with limited availability of labelled data. Additionally, out-of-sample validation tests evidenced the robustness and generalisability of the model when applied to pipelines not included during training. Overall, this study presents an effective and innovative solution for leak detection in water networks, reducing reliance on labelled data and improving system adaptability to real-world scenarios (Liu et al., 2020).

In 2024, Javier Ignacio Lobos Constanzo, under the supervision of Professor Yarko Niño Campos, presented his research at the University of Chile entitled *"Detection and Localisation of Leaks in Potable Water Distribution Networks in a Large City in Chile Using a Neural Network Classification Algorithm"*. The study employed a deep learning approach using neural networks, applying simulated data generated through EPANET. Scenarios with varying hydraulic conditions and leak locations were developed, focusing primarily on pressure variations as the main input variable.

The trained model achieved high classification accuracy in distinguishing between zones with and without leaks, demonstrating that it constitutes an effective tool for detecting and locating leaks in complex urban contexts. Although based on simulations, the approach is scalable and adaptable to real data, and represents a significant contribution towards improving preventive management and monitoring of water distribution networks (Constanzo & Ignacio, 2022).

In 2018, Juan Apesteguía Infantes and Edwin Huarcaya from the National University of Callao (Peru) developed a system based on artificial neural networks, using input variables such as flow and pressure, to recognise patterns associated with non-visible leaks in Lima's potable water network. Employing real flow and pressure data, the model trained in MATLAB successfully classified leak points with high accuracy, surpassing traditional manual methods. Its implementation enables real-time detection, reduces non-revenue water, and enhances the efficiency of urban water resource management (Apesteguía & Huarcaya, 2018).

### 2.3 Hybrid and Physics-Based Models

These models combine hydraulic equations (e.g., Hazen–Williams and Darcy–Weisbach) with artificial intelligence algorithms to generate relevant features and improve the localisation and quantification of leaks.

In 2020, Zahra Fereidooni and Hooman Tahayori from Shiraz University (Iran), alongside Ali Bahadori-Jahromi from the University of West London (United Kingdom), published an article entitled *"A Hybrid Model-Based Method for Leak Detection in Large-Scale Water Distribution Networks"*. In this study, the authors proposed a hybrid methodology that integrates traditional hydraulic models (such as the Hazen–Williams and Darcy–Weisbach equations) with artificial intelligence algorithms, including decision trees, k-nearest neighbours (KNN), random forests, and Bayesian networks. The implementation of cost-effective flow sensors at pipeline junctions enables real-time data collection, facilitating the accurate detection and localisation of leaks—even in complex scenarios involving multiple failures. This approach offers an efficient and cost-effective solution for monitoring large-scale water networks (Fereidooni et al., 2020).

### 2.4 Logic-Based and Anomaly Detection Algorithms

Logic-based methods detect simple anomalies, whereas algorithms such as the Local Outlier Factor (LOF) identify complex and adaptive patterns in historical and real-time data.

In 2024, Kiran Joseph, Jyoti Shetty, Ashok K. Sharma, Rudi van Staden, P. L. P. Wasantha, Sharna Small, and Nathan Bennett from Victoria University and Greater Western Water in Melbourne, Australia, published the article *"Leak and Burst Detection in Water Distribution Network Using Logic- and Machine Learning-Based Approaches"*. This study proposed a hybrid approach combining logical rules and machine learning models to detect leaks and bursts in water distribution networks.

Historical and real-time data were obtained via SCADA sensors installed in a pumping network in the Sunbury region of Victoria. Logic-based algorithms were applied to detect evident anomalies using predefined rules, such as abrupt pressure drops. To address more complex and variable patterns, several machine learning models were integrated, including K-Means, DBSCAN, Isolation Forest, One-Class SVM, and Local Outlier Factor (LOF).

The LOF model demonstrated high efficacy in identifying subtle deviations that could indicate leaks or bursts, outperforming traditional methods. As part of the system development, a web platform was created integrating these models, enabling real-time monitoring and automatic alert generation, thereby facilitating rapid response and significantly enhancing the operational management of the system. This approach represents a robust, scalable, and accurate solution for optimising the management of water resources in urban networks (Joseph et al., 2024).

## 3 Experimentation

The experimental phase of this research was designed to evaluate the feasibility, accuracy, and operational integration of a perceptron-based model for detecting silent leaks in domestic toilet systems. In contrast to many advanced detection frameworks, which often require substantial computational resources, specialised infrastructure, or high-cost deployment, the present approach prioritises simplicity, low computational overhead, and ease of integration within household environments. The methodology sought to ensure technical reliability while maintaining practical applicability, enabling seamless incorporation into an Internet of Things (IoT) architecture for real-time monitoring and user alerting.

Experimental trials were conducted under realistic domestic conditions, closely replicating typical water consumption patterns. The YF-B10 Hall-effect flow sensor was selected for its high temporal resolution, allowing the precise measurement of key variables: total filling time and the total number of pulses per event.

These parameters were identified as strong indicators of anomalous flow behaviour associated with leaks. Data were manually labelled for supervised learning, facilitating binary classification between normal consumption and leakage. This design aimed to validate not only the algorithmic performance in controlled evaluation but also its practical relevance for domestic water conservation applications.

### 3.1 Acoustic Signal Processing

Time-domain, frequency-domain, and correlation-domain analysis techniques are employed to extract features from acoustic signals captured by sensors, accelerometers, or hydrophones, enabling both real-time and non-real-time detection.

In June 2022, researchers Harris Fan, Salman Tariq, and Tarek Zayed from the Department of Building and Real Estate at The Hong Kong Polytechnic University published the article *"Acoustic Leak Detection Approaches for Water Pipelines"* in the journal *Automation in Construction*.

This study presents a comprehensive review of acoustic techniques employed for leak detection in water distribution systems. The authors categorise acoustic detection methods into two principal types: real-time and non-real-time approaches. Devices utilised for acoustic signal acquisition are highlighted, including noise loggers, wireless sensor networks, accelerometers, hydrophones, and fibre optic sensors.

With regard to signal processing, the study explores various feature extraction techniques, including time-domain analysis, frequency-domain analysis, time–frequency analysis, and correlation analysis.

These methods enable the identification of acoustic signals generated by leaks and their differentiation from other sources of noise.

The article also addresses current challenges in acoustic leak detection, such as minimising false alarms and detecting minor or weak leaks. It underscores the necessity of improving the precision and reliability of detection systems, particularly in complex urban environments (Fan et al., 2022).

In August 2022, researchers Rangsarit Vanijjirattikhan, Sunisa Khomsay, Nathavuth Kitbutrawat and collaborators, affiliated with the Metropolitan Waterworks Authority of Thailand, published the article *"AI-Based Acoustic Leak Detection in Water Distribution Systems"* in the journal *Results in Engineering*.

The study presents a leak detection system that utilises artificial intelligence (AI) to analyse leak sounds collected by acoustic sensors within water distribution networks.

Acoustic data are stored and managed in the cloud, facilitating the training of machine learning models such as deep neural networks (DNN), convolutional neural networks (CNN), and support vector machines (SVM). Among these, the DNN model demonstrated superior performance in identifying leaks, even with a less complex architecture than that of the CNN.

Furthermore, the system includes a mobile application that guides operators to the precise location of leaks, thus facilitating its use even by personnel with limited technical experience.

This approach enhances efficiency and accuracy in leak detection, contributing to water conservation and the reduction of losses within distribution networks (Vanijjirattikhan et al., 2022).

ISSN 2007-9737



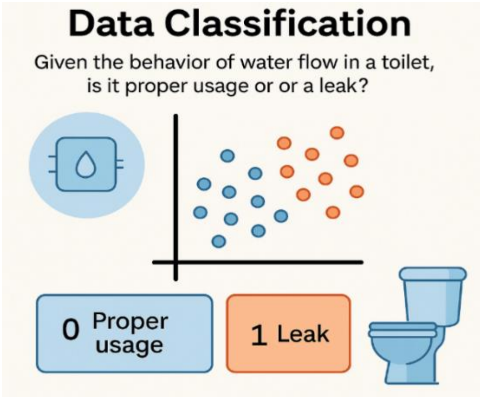


**Fig. 1.** Data classification for water leak detection

### 3.2 Graph Partitioning and Sensor Optimisation Algorithms

Graph partitioning algorithms and sensor placement optimisation techniques minimise the number of required measurements and enhance efficiency in leak identification.

In September 2021, researchers Zukang Hu, Wenlong Chen, Beqing Chen, Debao Tan, Yu Zhang, and Dingtao Shen, affiliated with Hohai University and the Changjiang Scientific Research Institute in China, published the article *"Robust Hierarchical Sensor Optimisation Placement Method for Leak Detection in Water Distribution System"* in the journal *Water Resources Management*.

This study proposed a robust hierarchical optimisation method for sensor placement in water distribution systems aimed at improving leak detection capabilities. The approach is based on sequential sensor selection using joint entropy as the objective function, while considering individual sensor failure scenarios to minimise information loss in the event of malfunction.

The methodology was validated through simulations using the EPANET example network Net3. The results demonstrated that, in the event of a sensor failure, the proposed scheme reduced the joint entropy loss from 0.011 (in conventional schemes) to 0.007. This indicates greater robustness and adaptability of the method for the detection and identification of leaks.

In summary, the study presents an effective and resilient strategy for sensor placement in water distribution networks, improving detection accuracy and system resilience to sensor failures (Hu et al., 2021).

### 3.3 Data Classification

Data classification is a fundamental task in supervised learning that consists of assigning a label or class to a set of input data according to its similarity with previously known examples. In the context of this research, the objective is to classify the behaviour of water flow in a toilet as either *"adequate consumption"* (class 0) or *"leak"* (class 1), based on real data recorded through sensors (Fig. 1).

### 3.4 Simple Perceptron as a Binary Classifier

The simple perceptron is a neural network model proposed by Frank Rosenblatt in 1958, suitable for linearly separable binary classification problems. Its functioning is based on computing a linear combination of inputs and applying an activation function (step function) that determines the output class. Due to its low computational cost, ease of implementation, and capacity for supervised learning, it is particularly useful for practical applications such as leak detection.

### 3.5 Mathematical Structure of the Perceptron

A simple perceptron comprises the following components (Fig. 2):

– **Inputs $(x_1, x_2)$**: In this case, two input variables representing the behaviour of water flow:

  $x_1$ **:** Total filling time of the toilet, measured in milliseconds.

  $x_2$: Total number of pulses accumulated by the sensor during filling.

– **Synaptic weights $(w_1, w_2)$**: Adjustable parameters that determine the relative importance of each input.

– **Bias $(b)$**: An additional value that allows adjustment of the decision boundary independently of the input values.

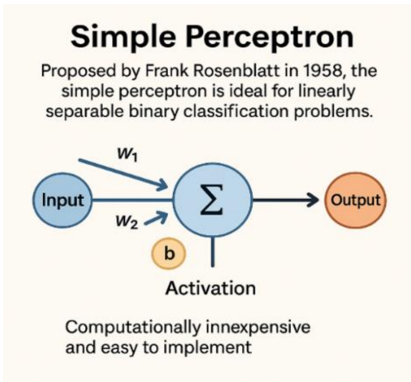


**Fig. 2.** Integration and structure of the simple perceptron

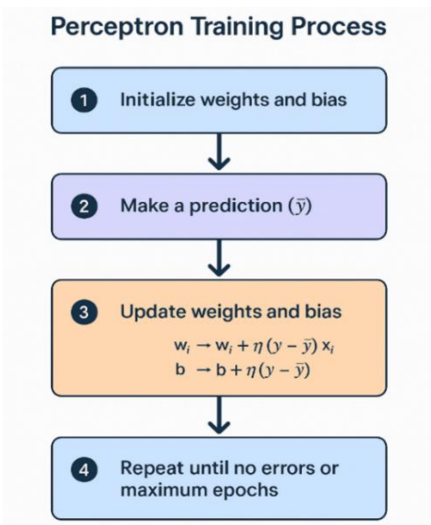


**Fig. 3.** Algorithm training using the perceptron learning rule

- **Linear combination** ($z$): $z = w_1 x_1 + w_2 x_2\, b$.
- **Activation function**: The binary step function is used:

$$f(z) = \begin{cases} 1 \;\; si \; z \geq 0 \\ 0 \;\; si \; z \leq 0 \end{cases}.$$

- **Output** ($\hat{y}$): Prediction result, taking the value 1 (leak) or 0 (adequate consumption)

### 3.6 Perceptron Training Process

Training the perceptron involves adjusting the weights and bias through a supervised learning algorithm that minimizes classification errors using the perceptron learning rule:

$$w_i \coloneqq w_i + \eta(y - \hat{y})x_i b \coloneqq b + \eta(y - \hat{y}),$$

where

- $y$ actual class of the sample (0 o 1)
- $\hat{y}$ predicted class by the model
- $x_i$value of the $i$-th input variable
- $\eta$ learning rate, which controls the adjustment magnitude in each iteration (e.g., 0.01)

This procedure is repeated for each training sample over multiple epochs, enabling the model to find a decision boundary that correctly separates the classes (Fig. 3).

### 3.7 Dataset Acquisition and Construction

To train the model, a dataset was designed from experimental trials conducted under real conditions. Samples were collected corresponding to complete toilet tank filling events, using the YF-B10 sensor that records the number of pulses at 0.256 millisecond intervals Fig. 4).

For each data file:

- The **total filling time** was calculated by multiplying the number of samples by the time interval.
- The **total number of pulses** per sample was computed.
- Samples were manually labelled as class 0 (adequate filling) or class 1 (leak), depending on whether flow was interrupted or continuous.

The final dataset was structured with three columns: Total_time_ms, Total_Pulses and Class, enabling the perceptron to be trained with real-world data.

### 3.8 Data Normalization

Before training, the input variables are scaled to ensure they share a common magnitude. Z-score scaling is applied, defined as follows:

ISSN 2007-9737



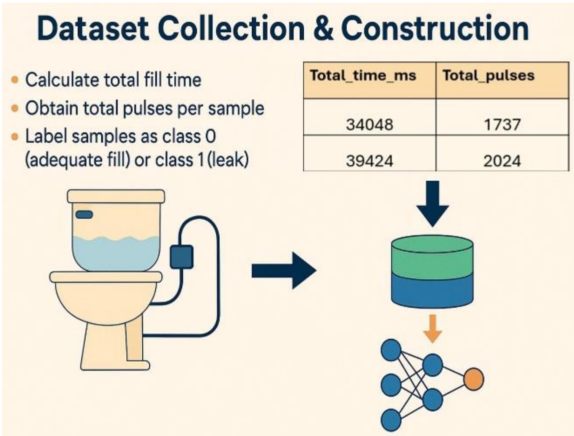


**Fig. 4.** Data acquisition and structure through real-use trials in toilets

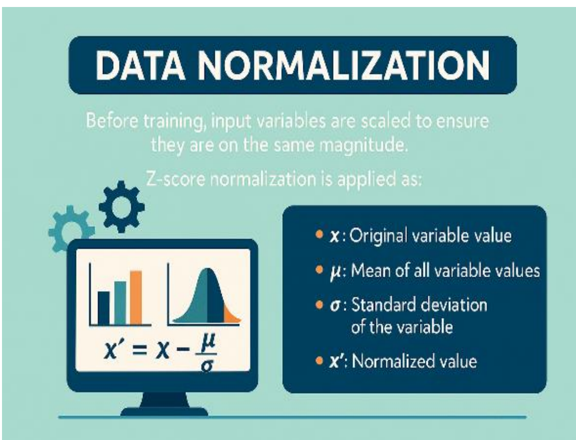


**Fig. 5.** Standardization of input variables through Z-score normalization

$$x' = \frac{x - \mu}{\sigma}$$

where:

$x$ : Original variable value,
$\mu$ : Mean of all values of the variable,
$\sigma$: Standard deviation the variable,
$x'$: Normalized value.

This procedure ensures that all variables have zero mean and unit variance, thereby enhancing the efficiency and stability of the training process (Fig. 5).

## 4 Results

The results obtained from the experimental implementation of the proposed perceptron-based detection system provide a comprehensive evaluation of its effectiveness in identifying silent toilet leaks under real domestic conditions. Performance metrics were derived from an independent test dataset, enabling an objective assessment of classification accuracy and error distribution. Particular emphasis was placed on the system's ability to correctly discriminate between adequate water consumption patterns and anomalous flow events indicative of leakage.

The integration of the algorithm within an Internet of Things (IoT) framework further allowed for continuous, real-time monitoring and instant user alerts, demonstrating the practical applicability of the approach beyond theoretical modelling. The findings presented in this section therefore address both the quantitative performance of the model—through measures such as accuracy and confusion matrix analysis—and its operational reliability within a functional monitoring system, offering valuable insights into its potential for large-scale residential deployment.

### 4.1 Interpretation of Results and Application

Once trained, the perceptron evaluates new samples using the linear combination of scaled inputs. If the result is greater than or equal to zero, a leak is inferred ($\hat{y}$ = 1); if it is less, it is classified as normal consumption ($\hat{y} = 0$).

For instance, if the model receives an input comprising a filling time and pulse count that yields a value of $\boldsymbol{z}$ = 0.45, the step function will classify the sample as a leak, since $\boldsymbol{z} \geq \mathbf{0}$. Regardless of whether the value is close to the threshold, the decision is definitive, which is ideal for alert systems.

### 4.2. Evaluation of Model Performance

The model's performance is assessed using metrics calculated over an independent test set. The following performance indicators were employed:

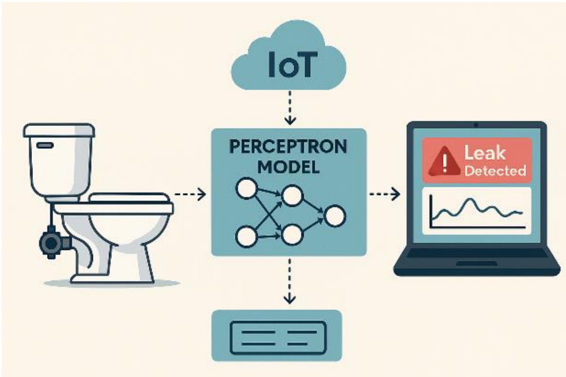


**Fig. 6.** Integrated architecture for detecting toilet water leaks via IoT and the simple perceptron

- **Accuracy:** The proportion of correctly classified samples relative to the total.
- **Confusion Matrix:** Displays the counts of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

In this study, the model achieved an accuracy of 98% on the test data, correctly classifying the samples corresponding to normal water consumption. It is recommended to balance the dataset with additional leak samples to obtain more representative performance metrics.

### 4.3 Real-time User Alerts

The simple perceptron represents a powerful tool for binary classification tasks involving real data. In the context of this project, its application was complemented by an Internet of Things (IoT) architecture, whereby flow sensor data were captured and processed in real time. This data acquisition structure included YF-B10 Hall-effect electromagnetic sensors, which recorded flow events at 0.256 millisecond intervals, allowing for high-precision readings.

Subsequently, the integration of the model with a graphical interface connected to an IoT system enabled real-time visualization of water flow behavior (Fig. 6). This system not only facilitated the automatic collection and analysis of data but also enhanced user responsiveness by generating immediate alerts when silent leaks or anomalous conditions were detected.

## 5 Conclusions

From a functional perspective, this solution promotes responsible water consumption, reduces economic losses caused by undesirable leaks, and contributes to operational sustainability in domestic installations. In contrast with previous research employing passive sensors or unsupervised learning strategies without real-time integration, this project offers a comprehensive technological proposal, encompassing detection and user-oriented decision-making visualization.

Ultimately, the performance achieved by the simple perceptron—demonstrated by an accuracy of 97% on test data—validates its suitability for this type of problem. Its low computational cost, interpretability, and compatibility with embedded systems confer a distinct advantage over more complex models.

This work presents a significant contribution to the digital era when compared with other studies, as it successfully integrates lightweight artificial intelligence, IoT monitoring, and immediate communication with the end user, thereby closing the full cycle of detection, analysis, and action.

## References

1. **Alves Coelho, J., Glória, A., Sebastião, P. (2020).** Precise Water Leak Detection Using Machine Learning and Real-Time Sensor Data. IoT, Vol. 1, No. 2, pp. 474–493. Doi: 10.3390/iot1020026.
2. **Apesteguía, J. A., Huarcaya, E. (2018).** Redes neuronales artificiales aplicadas a la detección de fugas no visibles en las redes de agua potable de la Ciudad de Lima.
3. **Choi, J., Sungbin I. (2023).** Application of CNN Models to Detect and Classify Leakages in Water Pipelines Using Magnitude Spectra of Vibration Sound. ResearchGate. Doi: 10.3390/app13052845.
4. **Constanzo, L., Ignacio, J. (2022)**. Detección y localización de fugas en redes de distribución de agua potable en una gran ciudad de Chile mediante un algoritmo de clasificación de redes neuronales. https://repositorio.uchile.cl/handle/2250/185782?utm_source.

5. **Fan, H., Tariq, S., Zayed, T. (2022).** Acoustic leak detection approaches for water pipelines. Automation in Construction, Vol. 138, 104226. Doi: 10.1016/j.autcon.2022.104226.
6. **Fereidooni, Z., Tahayori, H., Bahadori-Jahromi, A. (2020).** A hybrid model-based method for leak detection in large scale water distribution networks. Journal of Ambient Intelligence and Humanized Computing, Vol. 12. No. 2, pp. 1613–1629. Doi: 10.1007/s 12652-020-02233-2.
7. **Guo, G., Yu, X., Liu, S., Ma, Z., Wu, Y., Xu, X., Wang, X., Smith, K., Wu, X. (2021)**. Leakage Detection in Water Distribution Systems Based on Time–Frequency Convolutional Neural Network. Journal of Water Resources Planning and Management, Vol. 147, No. 2, 04020101. Doi: 10.1061/(ASCE)WR.1943-5452.0001317.
8. **Hu, Z., Chen, W., Chen, B., Tan, D., Zhang, Y., Shen, D. (2021).** Robust Hierarchical Sensor Optimization Placement Method for Leak Detection in Water Distribution System. Water Resources Management, Vol. 35, No. 12, pp. 3995–4008. Doi: 10.1007/s11269-021-02922-3
9. **Joseph, K., Shetty, J., Sharma, A. K., Van Staden, R., Wasantha, P. L. P., Small, S., Bennett, N. (2024)**. Leak and Burst Detection in Water Distribution Network Using Logic- and Machine Learning-Based Approaches. Water, Vol. 16, No. 14, pp. 1935. Doi: 10.3390/ w16141935.
10. **Lee, S., Kim, B. (2023).** Machine Learning Model for Leak Detection Using Water Pipeline Vibration Sensor. Sensors, Vol. 23, No. 21, Article 21. Doi: 10.3390/s23218935.
11. **Liu, S., Yu, J., Li, H., Wang, K., Wu, G., Wang, B., Liu, M., Zhang, Y., Wang, P., Zhang, J., Wu, J., Jing, Y., Li, F., Zhang, M. (2020).** Controllable Drug Release Behavior of Polylactic Acid (PLA) Surgical Suture Coating with Ciprofloxacin (CPFX)—Polycaprolactone (PCL)/Polyglycolide (PGA). Polymers, Vol. 12, No. 2, pp. 288. Doi: 10.3390/polym12020288.
12. **Peng, H., Xu, Z., Huang, Q., Qi, L., Wang, H. (2024).** Leakage Detection in Water Distribution Systems Based on Logarithmic Spectrogram CNN for Continuous Monitoring. Journal of Water Resources Planning and Management, Vol. 150, No. 6, 04024015. Doi: 10.1061/ JWRMD5.WRENG-6276.
13. **Ravichandran, T., Gavahi, K., Ponnambalam, K., Burtea, V., Mousavi, S.J. (2021)**. Ensemble-based machine learning approach for improved leak detection in water mains. Journal of Hydroinformatics, Vol. 23, No. 2, pp. 307–323. Doi: 10.2166/hydro.2021.093.
14. **Shukla, H., Piratla, K.R. (2023).** Leakage Detection in Water Distribution Network Using Machine Learning. pp. 192–200. Doi: 10.1061/ 9780784485033.021.
15. **US EPA, O. (2024).** Fix a Leak Week https://www.epa.gov/watersense/fix-leak-week.
16. **Vanijjirattikhan, R., Khomsay, S., Kitbutrawat, N., Khomsay, K., Supakchukul, U., Udomsuk, S., Suwatthikul, J., Oumtrakul, N., Anusart, K. (2022).** AI-based acoustic leak detection in water distribution systems. Results in Engineering, Vol. 15, 100557. Doi: 10.1016/ j.rineng.2022.100557.
17. **Xylem (2025).** Non-revenue water in LATAM: The role of AI in cutting water losses | Xylem Saint Vincent and the Grenadines. https://www.xylem.com/en-vc/support/lets-solve-water-blog/non-revenue-water-in-latam/

*Article received on 09/01/2025; accepted on 24/07/2025.*
**Corresponding author is Francisco Rafael Trejo-Macotela.*