Journal of Physical Chemistry and Functional Materials

Home Page of Journal: https://dergipark.org.tr/jphcfum

Machine Learning Methods in IoT Based Embedded Systems for Classifying Physical Faults in Water Distribution Networks

Irfan Kilic[a]*, Orhan Yaman[b] , Şeyma Saylan[b], İlayda Hörgüşlüoğlu[b], Betül Demirelli[b]

[a]*Firat University, Department of Information Technologies, Elazig, Turkey*

[b]*Firat University, Department of Digital Forensic Engineering, Elazig, Turkey*

* *Corresponding author: E-mail: irfankilic@firat.edu.tr*

ABSTRACT

Water is the most important factor for the survival of living things on Earth. Although 70% of the Earth is water, the amount of drinkable water is approximately 0.3%. Therefore, creating a sustainable water policy and carrying out studies are very important for our world and our future. Most of the potable water resources are physical losses. In the evaluations made based on metropolitan municipalities, it was seen that the water loss rate was approximately 50%. The study aims to find water pipe faults using IoT (Internet of Things) based machine learning classifiers to prevent physical losses in water distribution networks. Within the scope of this study, an experimental environment was created and an IMU (Inertial Measurement Unit) sensor was fixed on plastic pipes of different diameters and lengths. Vibration data collected in different scenarios (pressure, etc. factors) were transferred to the ThingSpeak platform over the internet. The transferred data could be monitored in real-time on a server. Physical damage in the pipes was detected using signal pre-processing, feature extraction, and feature selection algorithms on vibration data. In the study, damages were classified using machine learning-based classification (Decision Trees, k-Nearest Neighbors, Linear Discriminant, Support Vector Machines) methods to predict the type of damage (solid, hole, multi-hole). The data set revealed within the scope of the study is thought to lead to scientific studies in this field. The results obtained are close to the state-of-the-art results.

ARTICLE INFO

Keywords:

Water pipe faults

IoT-based embedded systems

ESP

Fault classification

Thingspeak platform

Machine learning

Received: 2024-11-19

Accepted: 2024-11-26

ISSN: 2651-3080

DOI: 10.54565/jphcfum.1588037

1. INTRODUCTION

Water is the most important source of life to sustain life. The rapidly increasing global warming problem in recent years is disrupting the balance in the ecosystem. Access to drinking water is becoming more difficult day by day, especially due to the increasing population and unconscious use of water. It can be said that if this situation continues, it is inevitable that water wars will start shortly. It can be said that 40% of the world, approximately 80 countries, are currently suffering from water shortage. Water problems that may be experienced by the country in 2040 are presented in Figure 1. According to numerical data; If the amount of water per person falls below 1000 m^3, it is indicated as the beginning of water scarcity. For a country to be considered water-rich, the amount of water per capita must be 10000 m^3. When we look at the water situation of our country, it is seen that the amount per capita is 1500 m^3. In other words, our country may become a country that may experience water problems shortly. According to 2010 Turkish Statistical Institute (TSI) data, the total amount of lost and illegally used water is approximately 2.2 billion m^3. In the evaluations made based on metropolitan municipalities, the water loss rate is approximately 50%, when evaluated on a provincial basis; The water loss rate increases to 60% in municipalities with a population of 100.000 to 150.000, and 60 to 70% in municipalities with a population of 100.000 to less than 100.000 [1]. The non-revenue water amounts of some cities in our country in 2015 are shown in Figure 2.

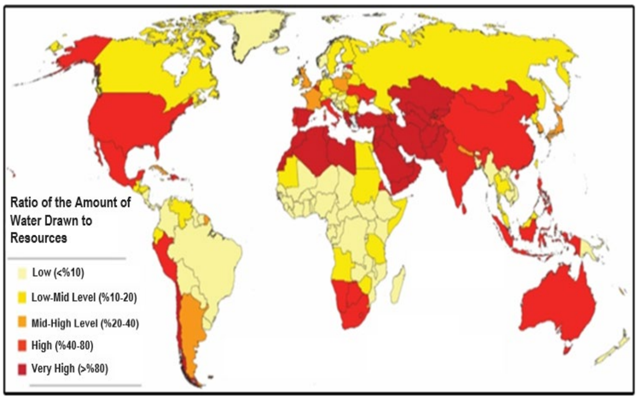


Figure 1 Water problems that may occur by country in 2040 [2]

Therefore, developing a sustainable water policy in the management of water resources and carrying out studies to prevent water waste is very important for our world, our future, and the continuation of life. Drinking water, which has a limited quota, is offered to people through infrastructures established in cities. Water losses are also increasing due to the deformation of the drinking water network infrastructure over the years. Due to limited potable water resources and increasing access to water and network costs, rapid detection of losses and leakages in networks has become important [4]. In drinking water, the amount of water entering the system is divided into permitted consumption and water loss. Water losses are called administrative and physical losses. Physical losses; It consists of losses and leakages occurring in supply and distribution lines and service connections, and leakages and overflows occurring in warehouses [3], [5]. The water balance structure used in the literature is presented in Figure 3 [3].

Literature Background

Within the scope of the study, studies on physical losses were focused on reducing the amount of water that does not generate income. Studies conducted in the literature to detect physical losses are summarized in Table 1. A general taxonomy of water intake systems is given in Figure 3

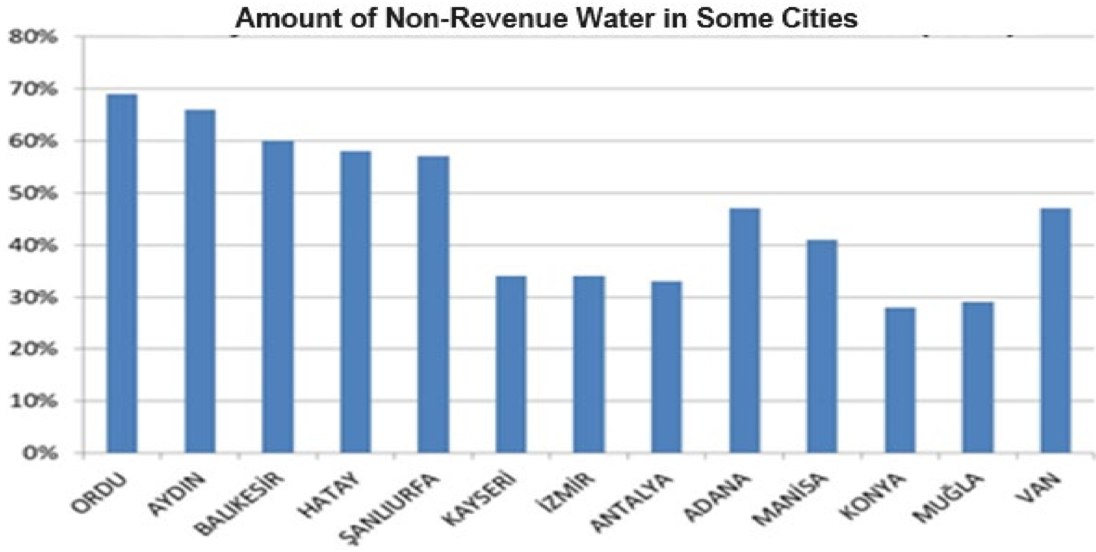


Figure 2 Non-revenue water amounts in some cities in our country in 2015 [3]

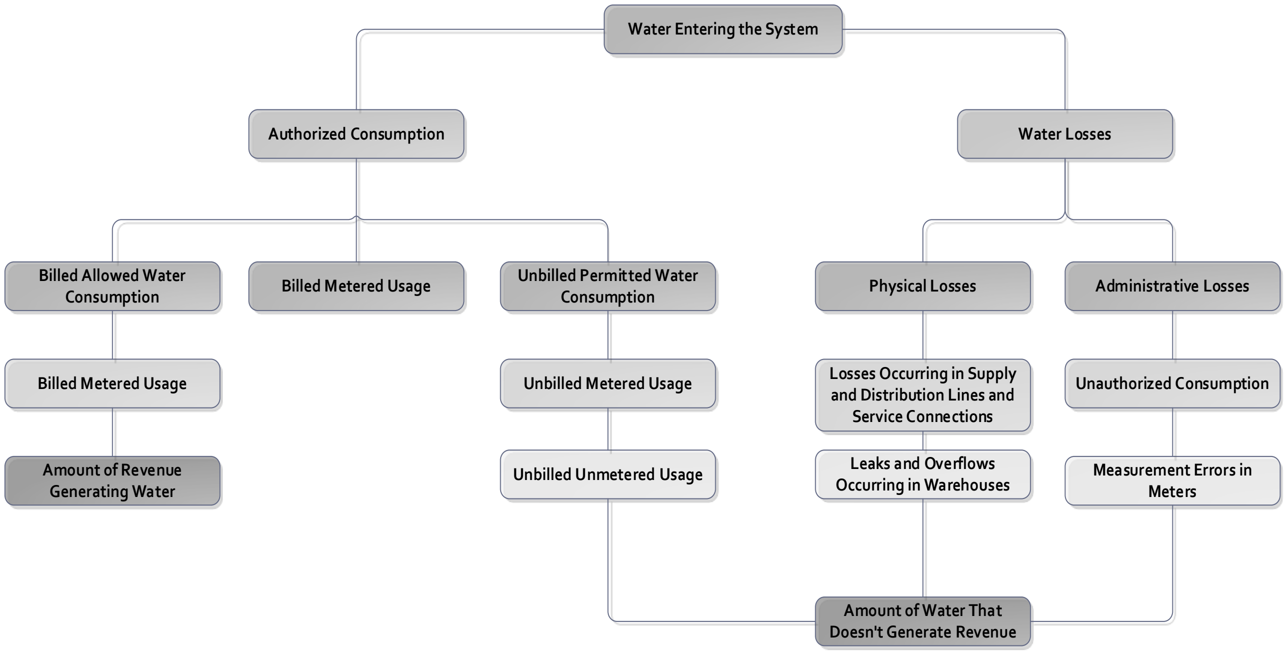


Figure 3 Water balance structure used in the literature [3]

Table 1 Studies in the literature to detect water physical losses

Reference, Year	Monitored Parameter	Sensor	Pipe Material	Pre-processing technique	Leak Detection and Recognition Algorithm
Kartakis et al. [6], 2016	Surface vibration of the pipe	Vibration Sensor (NEC Tokin)	-	Kalman Filtering	Compression Ratios Analysis
Ismail et al. [7], 2019	Surface vibration of pipe	Accelerometer (MPU6050, ADXL335 VE MMA7361)	Plastic (polietilen)	-	Offline Analysis
Karray et al. [8], 2016	Pressure	Force Sensitive Resistor	Plastic (polietilen)	Kalman Filtering and Compression	Estimated Kalman Filter
Stoianov et al. [9], 2007	Acoustic Signals and Pipe Surface Vibration	Hydrophones and Accelerometers	Plastic (Polivinil Klorür)	Fast Fourier Transform (FFT) and Compression	Acoustic Leak Detection Technique
Sadeghioon et al. [10], 2014	Pressure (Force Sensitive Resistor)	Temperature and Pressure	Plastic (Polivnil Klorür)	-	Relative Pressure Change
Nicola et al. [11], 2018	Acoustic Signals	Acoustic Sensors	Metallic	-	Acoustic Emission Technique
Marmarokopos et al.[12], 2018	Surface Vibration of Pipe	Accelerometer (KB12(VD))	Plastikc (polietilen)	Moving Average	Fast Fourier Transform, Wavelet Transform, Power Spectral Density, and Cra Spectral Density
Okosun et al. [13], 2019	Surface Vibration of Pipe	Piezoelectric transducer	Plastic (Polivinil Klorür)	Amplification	Amplitude Thresholding and FFT
Martini et al.	Surface	IEPE Accelerometer	Plastic	Signal Filtering and	Standard Deviation

[14], 2015	Vibration of Pipe		(Polietilen)	Amplification	
Choi et al. [15], 2017	Surface Vibration of Pipe	Vibration Sensor	-	-	Power Spectral Density and Coma Spectral Density
Songur et al. [16], 2021	Pressure and acoustic values	Flowmeter and interpolation	-	Detection of leakage location by interpolation on the network using noise intensity	
N. Yussof and H. Ho [17], 2022	Review	-	-	-	Review
Fan, X. et al. [18], 2022	Pipe physical values	Pipe-information dataset, Pipe maintenance dataset	-	Label assignment for dataset and data aggregation	LightGBM, ANN, Logistic Regression, K-NN, SVM
Şahin, E. and Yüce, H. [19], 2023	Pressure and Flow	Pressure sensors (MPS 500)	PVC Pipe	-	Graph Convolutional Neural Network, SVM

As can be seen in Table 1, as a monitoring parameter for the detection of physical losses;

- Surface vibration of the pipe,
- Pressure,
- Acoustic signals were used.

To monitor these parameters, sensors or sensor groups such as vibration sensors, accelerometers, force-sensitive resistors, hydrophores, temperature and pressure measuring sensors, acoustic sensors, piezoelectric transducers, and flow meters are used. These sensors can be used on metal and plastic pipes. When the studies in the literature are examined, it can be seen that there are almost no IoT-based systems for transferring sensor data. This situation constitutes the motivation for the project. Based on the studies carried out, the project has a structure that can be integrated into real life. Within the scope of the study, it was aimed to develop an IoT (Internet of Things) based system to detect losses in drinking water.

Study Motivation, Purpose, and Contributions

Drinking water in cities is classified as income-generating and non-income-generating water. The majority of the non-revenue water amount consists of physical losses. There are many reasons for physical losses in drinking water in cities. Some of these reasons are listed below;

- Most of the drinking and potable water networks have completed their service life,
- Inadequate and irregular maintenance and repairs,
- Occurrence of problems due to other infrastructure works,
- Employing technical personnel who do not have sufficient experience,
- Old infrastructure network made of cast/ductile iron or steel,
- Corrosion (inner + outer surface corrosion, material loss, pitting),
- Calcification, crusting
- Increased internal surface roughness,
- New malfunctions and new crack leaks as a result of increased pump pressure requirements,
- Can be summarized as earthquake and ground movements [20].

Sample images of underground faults that cause physical losses in drinking water are shown in Figure 4.

Figure 4 Example images of underground faults causing physical losses in drinking water

To use our water resources most effectively, it is important to develop methods and technologies to detect and reduce loss-leakage rates. Activities that can be done to minimize physical water losses are listed below:

- Infrastructure and asset management,
- Repair speed and quality,
- Pipeline management,
- Active leak control,
- Can be summarized as pressure management [21].

The main purpose of the study:

- Detection of water losses caused by physical damage,
- Initiating work to repair detected damages,
- Supporting infrastructure works,
- It is the most efficient use of potable water resources.

The scientific objectives of the study are;

- Establishing an experimental environment in a 50 cm deep and 20 m long channel with plastic water pipes of different pipe diameters for possible situations that may arise,
- Creating different scenarios through the established experimental environment,
- Transferring data collected using IoT devices to the ThingSpeak platform in real-time,
- Parsing and evaluating the data stored on the ThingSpeak platform with feature extraction and feature selection algorithms,
- It is the development of an artificial intelligence algorithm with at least 80% accuracy with the collected data.

In social terms, the study aims to remove society from the chaos that may occur in case of water shortage in the coming years, as a country suffering from water stress. Thus, it is anticipated that people's quality of life will increase.

The motivation and research questions required for the study and the contributions expected to be obtained as a result of the study are listed below.

Motivation and research questions;

- Can physical losses in mains water be detected using IMU sensors?
- Can data from sensor networks be transferred in real-time (at least 100 data per second for each sensor node) using IoT technology?
- Can the location of physical damage in the pipe be determined from vibration data?
- Can the extent of physical damage to the pipe be determined from vibration data?

Contributions;

- Real-time detection of physical damage in plastic pipes using vibration data,
- Removing noise in vibration signals using noise reduction algorithms,
- Transferring the signals obtained from ESP sensor nodes to the ThingSpeak platform in real-time,
- Using feature extraction and feature selection algorithms from vibration data,
- Physical fault detection in plastic pipes using deep learning and machine learning methods.

2. MATERIAL AND METHOD

Many problems are solved using IoT applications within the scope of smart cities. There are many IoT applications, especially for smart municipalities. In recent years, IoT-Cloud technology infrastructure has begun to be established within the scope of smart cities. For this reason, it is aimed to actively use IoT platforms in the project. In this study, IoT solutions are developed to detect water losses due to damage to water pipes while distributing mains water. Mains water is distributed to city centers through underground pipes. These pipes may be subject to wear and damage underground for many reasons. For example, if a small piece of stone touches the pipe surface at regular intervals, it is predicted that there will be wear on the contact surface and water leaks may occur. Due to many factors, the underground water distribution system can be damaged and if not detected early, it can lead to serious water loss. There are many studies in this field that we mentioned in the introduction section For example, there are studies such as the detection of pressure changes in water pipes or monitoring with an underwater robot placed inside the pipe [8], [10].

Data Collection

The IoT-based data collection platform developed within the scope of the study is presented in Figure 5.

As can be seen in Figure 5, in the proposed method, the sensor node (Vibration meter and ESP module) is placed on the pipes. Thanks to the WiFi feature on this sensor module, it is connected to the nearest center (4.5G internet module). ESP modules connected to the central nodes via WiFi are set to send vibration (x, y, and z axes) data. The data coming to the central node is transferred to the Thingspeak platform with 4.5G internet modules. There is a SIM card in the module to be used as a 4.5G module. By sharing SIM card data, a WiFi area network is created and sensor nodes can be connected. ThingSpeak is a freely available data platform designed for IoT applications. It provides storage and monitoring of real-time data. It also enables the collected data to be analyzed with MATLAB [22], [23].

Our data set, which we obtained with the help of the data collection platform in Figure 5, is given in Table 2.

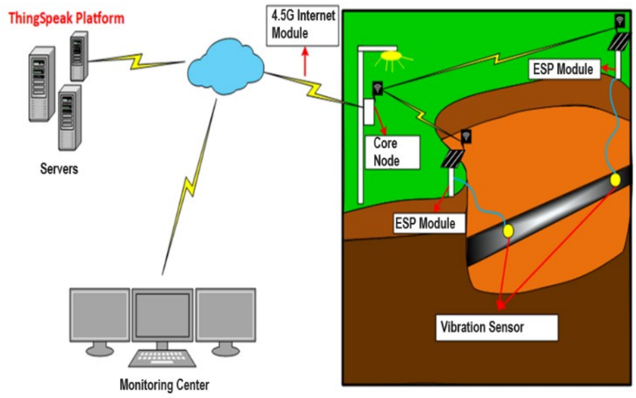


Figure 5 IoT-based data collection platform developed within the scope of the study

Table 2 Our data set is class and data numbers

Number	Class	Number of Data
1	Solid	25.897
2	Hole	25.698
3	Many Hole	26.258

Proposed Method

The general connection of the sensor node module developed within the scope of the study is shown in Figure 6.

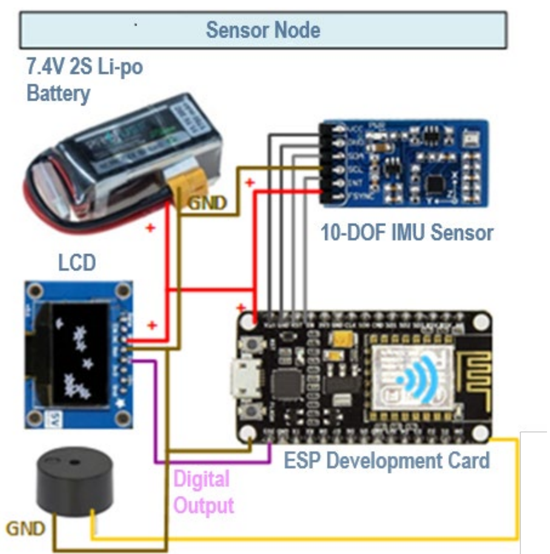


Figure 6 General connection of the sensor node module developed within the scope of the study

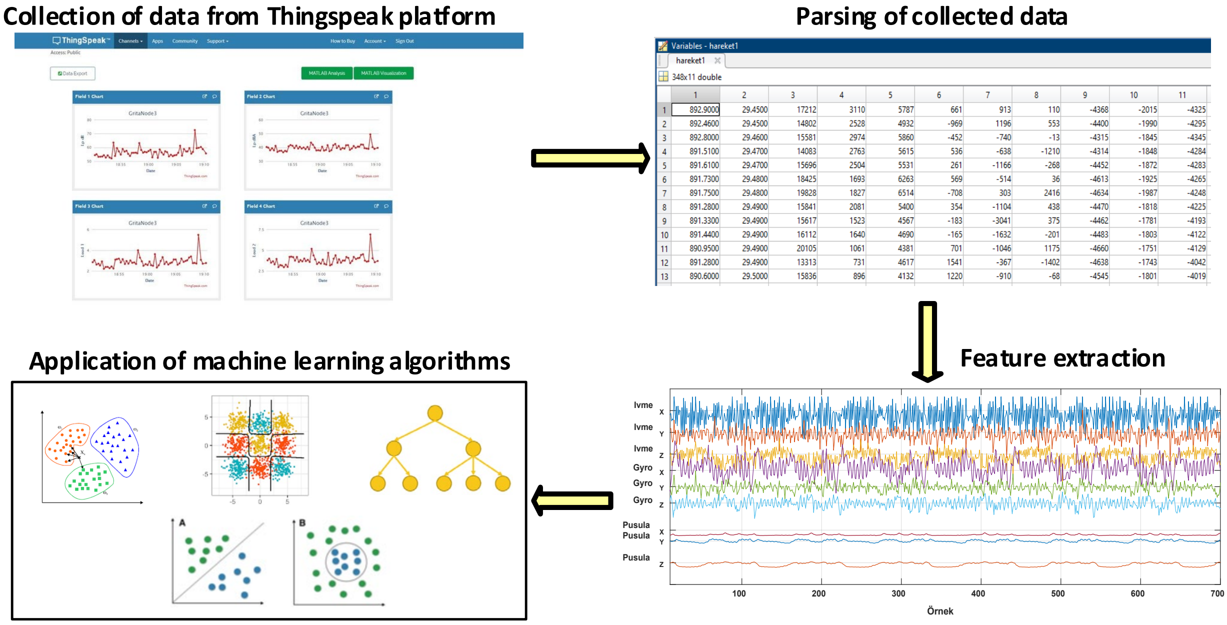


Figure 7 Block diagram of the method developed within the scope of the study

The first step in the method developed in the study was to create an IoT platform. Vibration data collected from the sensor nodes in the experimental setup was transferred to the ThingSpeak platform. In the proposed method, the data collected on this platform is transferred to the computer and parsed. Sensor IDs were taken into consideration during the parsing phase. By using multiple sensor nodes, the data collected was separated and feature extraction was made. The most weighted features are selected from the features obtained through feature extraction. The selected features were detected for leaks using machine learning-based methods. The results obtained were compared with the methods in the literature and the method was improved.

Decision Trees (DT)

Decision Trees (DA) are a method popularly used in operations research, particularly in decision analysis, to help determine the strategy most likely to achieve a goal, and in machine learning. A decision tree is a flowchart-like structure in which each internal node represents a "test" on an attribute (e.g., whether a coin toss lands heads or tails), each branch represents the outcome of the test, and each leaf node represents an attribute. The decision taken after all the features are calculated determines the class. Paths from root to leaf represent classification rules [24], [25].

A decision tree consists of three types of nodes.

- Decision nodes (□ square shaped)
- Chance nodes (○ circular)
- End nodes (Δ triangular shape)

k-Nearest Neighbors Algorithm (k-NN)

In statistics, it is a non-parametric trainer machine learning method, first developed in 1951 by E. Fix and J. Hodges [26] and later extended by T. Cover and P. Hart [27]. k-NN, used for classification and regression purposes, consists of the k-closest training data in an input data set. The output varies depending on whether k-NN is used for classification or regression [28].

Linear Discriminant Analysis (LDA)

Linear discriminant analysis (LDA is a generalization of Fisher's linear discriminant, a method used in statistics and machine learning to find a linear combination of features that characterize or separate two or more classes of objects or events. The resulting combination can be used for dimensionality reduction as a linear classifier.

LDA is closely related to the analysis of variance and regression analysis, which also attempts to express a dependent variable as a linear combination of other features or measures [29], [30]. LDA is also closely related to principal component analysis (PCA) and factor analysis; because they both look for linear variations of the variables that best explain the data. LDA specifically attempts to model the difference between data classes [31].

Support Vector Machines (SVM)

It is a machine-learning method used in classification problems. SVM determines a decision boundary between the two classes that are furthest from any point in the training data. This decision boundary gives a hyperplane that separates the two classes as best as possible. SVM represents the training data in a space called vector space.

A vector space is a coordinate system that represents each data point as a vector. The size of vectors depends on the number of features of the data. SVM uses it as a label that represents the class of each vector in the training data. For example, if data in a dataset is labeled as cat and dog, SVM tries to find a decision boundary to separate these data as belonging to a cat or dog class. SVM can use the hard margin or soft margin approach to find the decision boundary between two classes [32], [33], [34].

Experiment Setup

Within the scope of this study, an embedded system was designed using the ESP8266 development board and MPU6050 sensor. This embedded system aims to detect cracks in irrigation pipes. Vibration data was collected by fixing the embedded system board on the developed experimental setup. The algorithm pseudocode of the developed embedded system is given in Table 3.

Table 3 Algorithm pseudo-code running on the embedded card

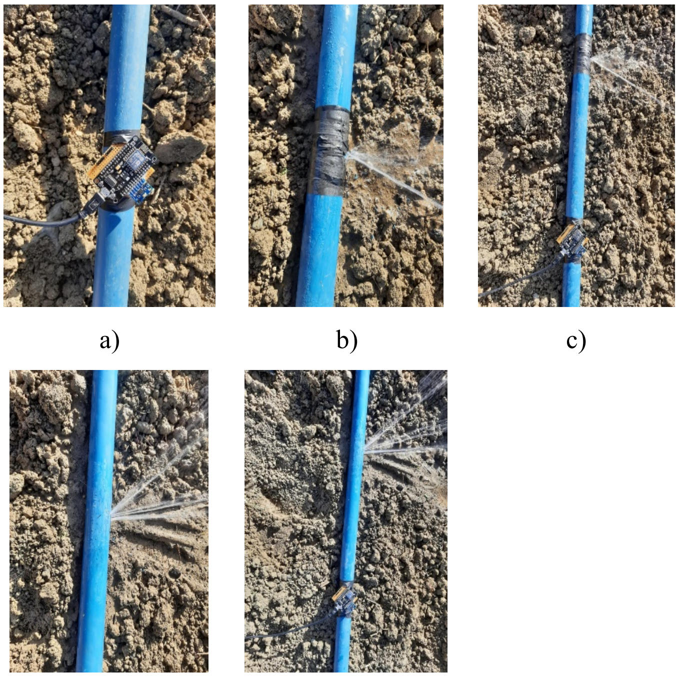
Step No	**Steps**
00	Load relevant libraries (adafruit_mpu6050, sensor, wire)
01	Start Program
02	Wait until the connection is established
03	Start serial connection
04	If no connection is established, wait and go to step 03
05	If the connection is established, set motion detection
06	Wait to receive sensor events
07	Fetch new sensor events with readings
08	Finish

Images of the embedded system board and experimental environment are given in Figure 8.

3. EXPERIMENTAL RESULTS DISCUSSION

Vibration data was collected from the experimental environment for 3 different situations. First of all, signals were collected from the embedded system for the intact pipeline. Then, signals were obtained from through-hole and multi-hole pipelines. Thus, a data set consisting of 3 classes was collected. The collected data set includes acceleration_x, acceleration_y, acceleration_z, gyro_x, gyro_y and gyro_z features. For this data set, fault detection was performed in the MATLAB version 2022b environment using machine learning algorithms, Decision Trees (KA), k-Nearest Neighbors (k-NN), Linear Discriminant Analysis (DDA), Support Vector Machines (SVM). The Scatter Plot chart created using MATLAB Classification Learner Toolbox is given in Figure 9.

Confusion matrices of the classifiers applied to the data set are presented in Figure 10.



a) b) c)

c) d)

Figure 8 Collecting data on solid and perforated pipes a) Placing the sensor module on the solid pipe b) Perforated pipe c) Placing the sensor module on the perforated pipe d) Multi-hole pipe e) Placing the sensor module on the multi-hole pipe

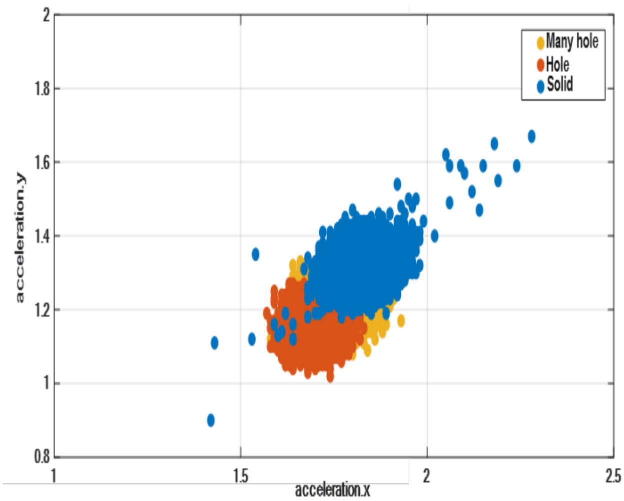


Figure 9 Scatter plot graphic

When the confusion matrices given in Figure 10 are examined, it is seen that the results obtained for 3 different classes are evenly distributed.

In Table 4, accuracy values are given using the class by class 4 classifier for 3 class types.

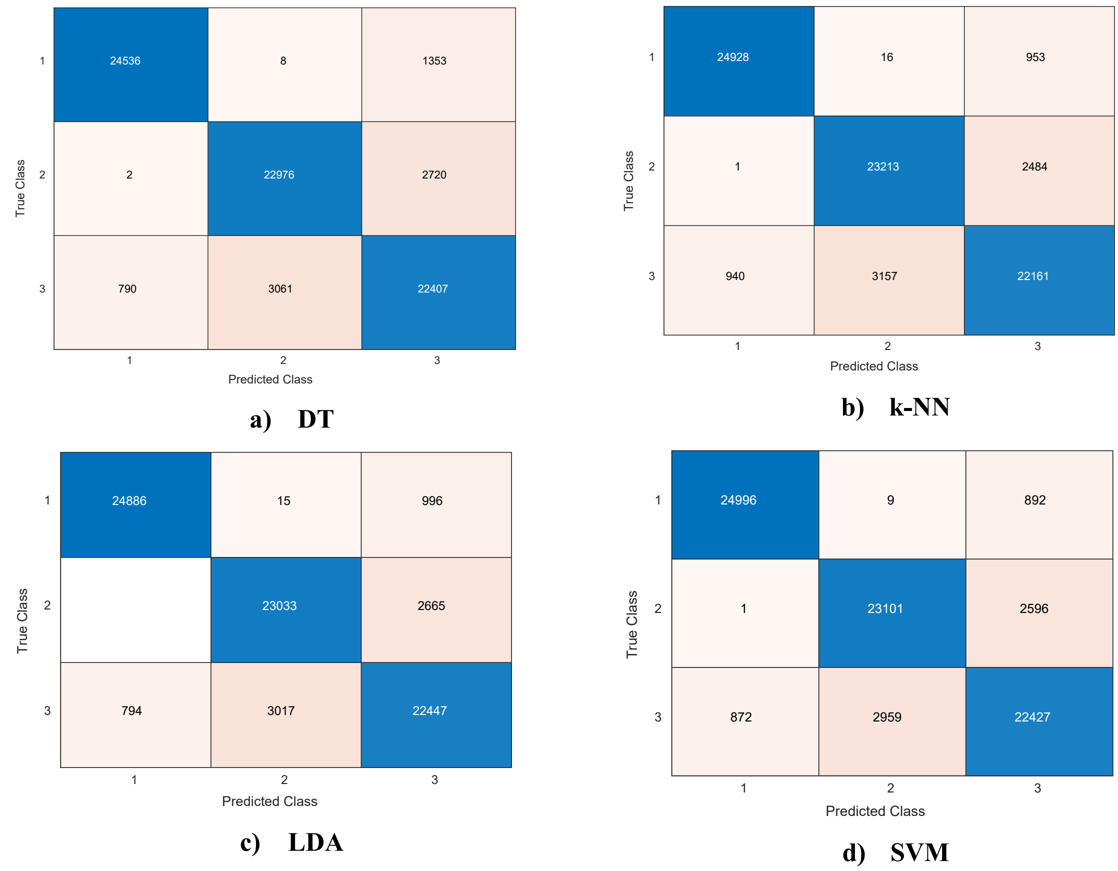


Figure 10 Confusion matrix results obtained in the proposed method

Table 4 Class-by-class accuracy values

Classifier	Solid (%)	Hole (%)	Many Hole (%)
DT	**96.10**	89.63	**85.47**
k-NN	**96.22**	**90.24**	84.45
LDA	94.83	89.29	**85.47**
SVM	**96.48**	89.87	85.40

When Table 4 is examined, it is seen that the accuracy results for the 3 class types are close to each other. It is seen that the k-NN classifier gives the best result (96.22%) for solid pipes. The results of other classifiers are also close to this. For hollow pipes, the best result (90.24%) was obtained with the k-NN classifier. For other classifiers, the result is slightly different. For multi-hole pipes, DT and LDA give the best results (85.47%). Table 5 gives the running times of the training data for each classifier. Here, it can be seen that Decision Trees (DT) and Linear Discriminant Analysis (LDA) classifiers work much faster.

Table 5 Classifier training data running times

Classifier	Training time (s)
DT	**0.56 s**
k-NN	5.02 s
LDA	**0.59 s**
SVMs	26.42 s

Table 6 gives the performance values of Accuracy, Precision, Recall, Geometric Mean, and F1-Score for 3 different classifiers.

When Table 6 is examined, SVM gives the best result for 100 iterations, albeit with a very small difference. However, other classifier results also have very similar values. Considering the duration of education, KA and DDA were preferred. Figure 11 shows the graph of the accuracy values obtained in the training of 4 classifiers in our data set in 10 parts.

Table 6 Performance results of the proposed methods for 100 iterations

		Accuracy (%)	Precision (%)	Recall (%)	Geometric Mean (%)	F1-Score (%)
DT	Max	90.41	90.45	90.43	90.33	90.44
	Min	90.34	90.38	90.37	90.26	90.37
	Mean	90.38	90.41	90.40	90.30	90.41
	Std	0.012	0.012	0.012	0.012	0.012
k-NN	Max	90.32	90.32	90.35	90.22	90.33
	Min	90.22	90.23	90.25	90.12	90.24
	Mean	90.27	90.27	90.30	90.17	90.29
	Std	0.019	0.019	0.019	0.019	0.019
LDA	Max	89.86	89.95	89.88	89.80	89.92
	Min	89.72	89.80	89.75	89.66	89.78
	Mean	89.79	89.88	89.81	89.73	89.85
	Std	0.03	0.032	0.032	0.033	0.032
SVM	Max	**90.58**	**90.59**	**90.60**	**90.49**	**90.59**
	Min	90.56	90.57	90.58	90.47	90.57
	Mean	90.57	90.58	90.59	90.48	90.58
	Std	0.007	0.007	0.007	0.007	0.007

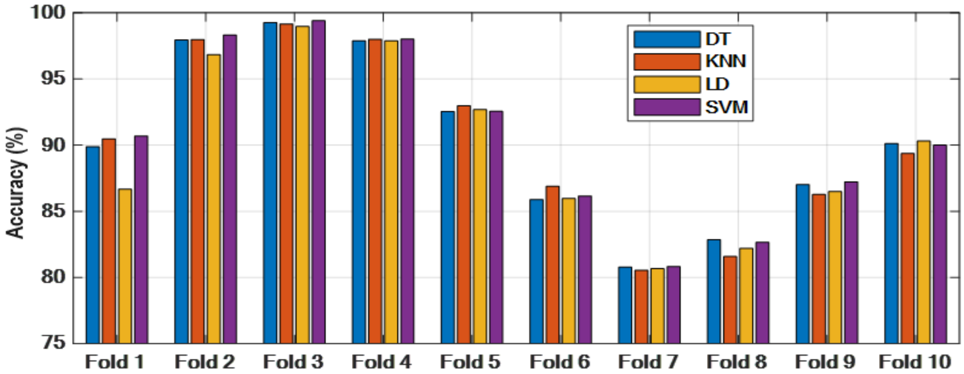


Figure 11 Accuracy values of classifiers for 10-piece training data set

4. CONCLUSION

Within the scope of the study, an experimental setup was created and a data set was collected to detect water leaks in water network lines. Using machine learning algorithms on the collected data set, solid, perforated, and multi-hole pipes were detected with an accuracy of over 90%. The results obtained are close to the state of the art compared to the size of our data set.

In future studies aim to develop lightweight algorithms with high accuracy. In addition, new methods will be developed by collecting data sets for underground pipelines.

Acknowledgement: This study was supported by TUBITAK 2209/A. Project number: 1919B012217701.

Competing interests

The authors declare that they have no competing interests.

REFERENCES

[1] T.C. Kalkınma Bakanlığı Özel İhtisas Komisyonu Raporu, "On Birinci Kalkınma Planı (2019-2023) Su Kaynakları Yönetimi ve Güvenliği," p. 110, 2018.

[2] D. K. Enstitüsü, "2040 Yılında En Çok Su Sıkıntısı Çekecek Ülkeler," 2022.

[3] Ç. C. Dilcan, G. Çapar, A. Korkmaz, Ö. İiritaş, Y. Karaaslan, and B. Selek, "İçme Suyu Şebekelerinde Görülen Su Kayıplarının Dünyada ve Ülkemizdeki Durumu," *Kalkınmada Anahtar Veriml. Dergisi, T.C. Sanayi ve Teknol. Bakanl. Aylık Yayın Organı*, vol. 30, no. 354, pp. 10–18, 2018.

[4] D. Karakaya and Z. F. Toprak, "İçme Suyu Şebekelerindeki Su Kayıplarının ZFT Classification Water Losses in Water Distribution Networks Using ZFT Algorithm," pp. 22–30, 2018.

[5] B. Ulusal and B. Raporuna, "İçme suyu dağitim sistemlerindeki kayiplar ve prowat projesi," pp. 1–8, 2007.

[6] S. Kartakis, W. Yu, R. Akhavan, and J. A. McCann, "Adaptive edge analytics for distributed networked control of water systems," *Proc. - 2016 IEEE 1st Int. Conf. Internet-of-Things Des. Implementation, IoTDI 2016*, pp. 72–82, 2016,

doi: 10.1109/IoTDI.2015.34

[7] M. I. Mohd Ismail *et al.*, "A review of vibration detection methods using accelerometer sensors for water pipeline leakage," *IEEE Access*, vol. 7, pp. 51965–51981, 2019, doi: 10.1109/ACCESS.2019.2896302

[8] F. Karray, A. Garcia-Ortiz, M. W. Jmal, A. M. Obeid, and M. Abid, "EARNPIPE: A Testbed for Smart Water Pipeline Monitoring Using Wireless Sensor Network," *Procedia Comput. Sci.*, vol. 96, pp. 285–294, 2016, doi: 10.1016/j.procs.2016.08.141

[9] I. Stoianov, L. Nachman, S. Madden, and T. Tokmouline, "PIPENETa wireless sensor network for pipeline monitoring," *IPSN 2007 Proc. Sixth Int. Symp. Inf. Process. Sens. Networks*, pp. 264–273, 2007, doi: 10.1145/1236360.1236396

[10] A. M. Sadeghioon, N. Metje, D. N. Chapman, and C. J. Anthony, "SmartPipes: Smart wireless sensor networks for leak detection in water pipelines," *J. Sens. Actuator Networks*, vol. 3, no. 1, pp. 64–78, 2014, doi: 10.3390/jsan3010064

[11] M. Nicola, C. Nicola, A. Vintilă, I. Hurezeanu, and M. Du, "Pipeline Leakage Detection by Means of Acoustic Emission Technique Using Cross-Correlation Function," *J. Mech. Eng. Autom.*, vol. 8, no. 2, pp. 59–67, 2018, doi: 10.5923/j.jmea.20180802.03

[12] K. Marmarokopos, D. Doukakis, G. Frantziskonis, and M. Avlonitis, "Leak detection in plastic water supply pipes with a high signal-to-noise ratio accelerometer," *Meas. Control (United Kingdom)*, vol. 51, no. 1–2, pp. 27–37, 2018, doi: 10.1177/0020294018758526

[13] F. Okosun, P. Cahill, B. Hazra, and V. Pakrashi, "Vibration-based leak detection and monitoring of water pipes using output-only piezoelectric sensors," *Eur. Phys. J. Spec. Top.*, vol. 228, no. 7, pp. 1659–1675, 2019, doi: 10.1140/epjst/e2019-800150-6

[14] A. Martini, M. Troncossi, and A. Rivola, "Automatic Leak Detection in Buried Plastic Pipes of Water Supply Networks by Means of Vibration Measurements," *Shock Vib.*, vol. 2015, pp. 11–15, 2015, doi: 10.1155/2015/165304

[15] J. Choi, J. Shin, C. Song, S. Han, and D. Il Park, "Leak detection and location of water pipes using vibration sensors and modified ML prefilter," *Sensors (Switzerland)*, vol. 17, no. 9, pp. 1–17, 2017, doi: 10.3390/s17092104

[16] M. SONGUR, A. DABANLI, B. YILMAZEL, and M. A. ŞENYEL KÜRKÇÜOĞLU, "Su Dağıtım Şebekelerindeki Fiziki Kayıpların Önlenmesinde SCADA'nın Önemi: ASKİ Örneği," *Afyon Kocatepe Univ. J. Sci. Eng.*, vol. 21, no. 6, pp. 1424–1433, 2021, doi: 10.35414/akufemubid.947662

[17] N. A. M. Yussof and H. W. Ho, "Review of Water Leak Detection Methods in Smart Building Applications," *Buildings*. 2022. doi: 10.3390/buildings12101535

[18] X. Fan, X. Wang, X. Zhang, and X. (Bill) Yu, "Machine learning based water pipe failure prediction: The effects of engineering, geology, climate and socio-economic factors," *Reliab. Eng. Syst. Saf.*, 2022, doi: 10.1016/j.ress.2021.108185

[19] E. Şahin and H. Yüce, "Prediction of Water Leakage in Pipeline Networks Using Graph Convolutional Network Method," *Appl. Sci.*, 2023, doi: 10.3390/app13137427

[20] M. Öztürk, "İçme suyu şebeke sistemi sanki halbur," *Independent Türkçe*, 2020.

[21] H. Muhammetoğlu and A. Muhammetoğlu, "İçme Suyu Temin ve Dağitim Sistemlerindeki Su Kayiplarinin Kontrolü," 2017, pp. 1-164.

[22] A. H. Miry and G. A. Aramice, "Water monitoring and analytic based ThingSpeak," *Int. J. Electr. Comput. Eng.*, vol. 10, no. 4, pp. 3588–3595, 2020, doi: 10.11591/ijece.v10i4.pp3588-3595

[23] H. Benyezza, M. Bouhedda, K. Dyellout, and A. Saidi, "Smart Irrigation System Based Thingspeak and Arduino," no. November, pp. 7–10, 2018.

[24] J. R. Quinlan, "Induction of Decision Trees," *Mach. Learn.*, 1986, doi: 10.1023/A:1022643204877

[25] B. Kamiński, M. Jakubczyk, and P. Szufel, "A framework for sensitivity analysis of decision trees," *Cent. Eur. J. Oper. Res.*, 2018, doi: 10.1007/s10100-017-0479-6

[26] E. Fix and J. L. Hodges, "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties," *Int. Stat. Rev. / Rev. Int. Stat.*, 1989, doi: 10.2307/1403797

[27] T. M. Cover and P. E. Hart, "Nearest Neighbor Pattern Classification," *IEEE Trans. Inf. Theory*, 1967, doi: 10.1109/TIT.1967.1053964

[28] M. Mohy-eddine, A. Guezzaz, S. Benkirane, and M. Azrour, "An Intrusion Detection Model using election-Based Feature Selection and K-NN," *Microprocess. Microsyst.*, p. 104966, 2023, doi: https://doi.org/10.1016/j.micpro.2023.104966. Available: https://www.sciencedirect.com/science/article/pii/S0141933123002107

[29] F. R. S. R. A. Fisher, Sc.D., "The use of multiple measurements in taxonomic problems," *Ann. Eugen.*, 1936.

[30] B. A. Moore and G. J. McLachlan, "Discriminant Analysis and Statistical Pattern Recognition.," *J. R. Stat. Soc. Ser. A (Statistics Soc.*, 1994, doi: 10.2307/2983518

[31] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2001, doi: 10.1109/34.908974

[32] C. Cortes and V. Vapnik, "Support-Vector Networks," *Mach. Learn.*, 1995, doi: 10.1023/A:1022627411411

[33] T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning, Second Edition," *Springer New York, NY*, 2009.

[34] S. Dhakshina Kumar, S. Esakkirajan, S. Bama, and B. Keerthiveena, "A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier," *Microprocess. Microsyst.*, 2020, doi: 10.1016/j.micpro.2020.103090