MACHINE LEARNING BASED SMART WATER LEAKAGE DETECTION LOCALIZATION AND CONSUMPTION FORECASTING SYSTEM

[1] G. Pranith B.Tech, M.Tech,(Ph.D) Assistant Professor, Department of CSE, Eluru College of Engineering and technology Duggirala (v), Pedavegi (m),Eluru-534004.
[2] Kalpukuri Manohar, M.Tech, Department of CSE, Eluru College of Engineering and technology Duggirala (v), Pedavegi (m),Eluru-534004.
[3]Gunaseharan Venugopal, Ph.D, MBA, Assistant Professor Department of Material Science School of Technology, Central University of Tamilnadu, Tiruvanur-610005 Tamilnadu.

Abstract: Water losses caused by hidden pipeline leakage create serious environmental, economic and operational challenges in domestic, institutional and municipal water-distribution networks.Since many pipelines are installed underground or in inaccessible locations, small leaks may continue for long periods before visible damage occurs. This project proposes a Machine Learning-Based Smart Water Leakage Detection, Localization and Consumption Forecasting System that combines Internet of Things sensing, real-time data processing and predictive analytics. Flow and pressure sensors installed at selected points continuously collect pipeline measurements and transmit them through an ESP32-based communication unit to a centralized monitoring platform. The collected time-series data are cleaned and transformed into features such as flow imbalance, pressure variation, rate of change, rolling averages and consumption lags. Machine-learning algorithms, including Random Forest, Support Vector Machine and Gradient Boosting, are evaluated to distinguish normal operating conditions from small, medium and major leakage events. The probable leakage location is identified at zone level by comparing pressure drops and inlet–outlet flow differences between neighboring sensor points. In addition to leakage management, the system forecasts future water consumption using historical usage patterns, time-based variables and optional environmental information. The forecasting module supports short-term demand planning, pump scheduling and efficient resource allocation. A web-based dashboard displays live sensor readings, leakage severity, probable leak zone, alerts, consumption trends and predicted demand. The proposed system provides a low-cost, explainable and scalable framework for reducing water loss, improving maintenance response and supporting sustainable water management.

Keywords: water leakage detection, leak localization, consumption forecasting, machine learning, Internet of Things, smart water management, flow sensor, pressure sensor, ESP32.

1. INTRODUCTION

Water is one of the most essential natural resources required for human life, agriculture, industries, and urban development. With rapid population growth, urbanization, and increasing demand for water, efficient management of water resources has become a major challenge. A significant amount of treated water is lost through pipeline leakage, damaged infrastructure, unauthorized consumption, and inefficient distribution systems. Conventional water management systems generally depend on manual inspection, periodic meter readings, and rule-based monitoring, which may not be sufficient for detecting leakages at an early stage or accurately predicting future water consumption. Therefore, intelligent and automated water management systems are increasingly important for reducing water losses and improving the efficiency of water distribution networks. Water leakage is particularly difficult to identify because leaks can occur at different locations and may remain undetected for long periods. Small leaks can gradually develop into major failures, resulting in substantial water wastage, infrastructure damage, and increased operational costs. Traditional leakage detection techniques often require physical inspection or specialized equipment, making them time-consuming and expensive, particularly for large and complex distribution networks. The availability of sensors and Internet of Things (IoT) technologies provides an opportunity to continuously collect information such as water flow rate, pressure, water consumption, and pipeline conditions. However, collecting sensor data alone is not sufficient; intelligent analytical techniques are required to identify abnormal patterns and determine the possible location of leakage. Machine Learning (ML) provides effective techniques for analyzing large volumes of water-network data and identifying patterns that may not be easily recognized using conventional approaches. Machine learning models can be trained using historical flow, pressure, consumption, and sensor data to distinguish between normal and abnormal water usage patterns. By analyzing deviations in these parameters, the proposed system can detect potential leakage conditions and estimate their probable locations. Depending on the available dataset, techniques such as Random Forest, Decision Tree, Support Vector Machine, K-Nearest Neighbors, Gradient Boosting, or Artificial Neural Networks can be employed for leakage classification and localization. The use of machine learning can improve the speed, accuracy, and automation of water leakage monitoring. In addition to leakage detection and localization, **water consumption forecasting** is an important component of efficient water resource management. Water demand varies according to factors such as time, season, weather conditions, population, day of the week, and consumption behavior. Accurate forecasting can help water authorities and consumers estimate future demand, plan distribution, identify unusual consumption, and optimize water supply. Machine learning-based forecasting models can learn historical consumption patterns and generate predictions for future water usage. Comparing predicted consumption with actual consumption can also provide an additional mechanism for identifying abnormal usage and possible leakage. The proposed **Machine Learning Based Smart Water Leakage Detection, Localization and Consumption Forecasting System** integrates these capabilities into a unified intelligent framework. Sensor or historical water-network data are first collected and preprocessed to remove noise, missing values, and inconsistencies. Relevant features such as flow rate, pressure, consumption rate, time, and historical usage patterns are then extracted. Machine learning algorithms are applied to classify water-network conditions as normal or potentially leaking, estimate the location or zone of the leakage, and forecast future water consumption. The system can provide alerts when abnormal conditions are detected, enabling timely maintenance and reducing unnecessary water loss. The proposed system can therefore support a transition from **reactive water management to predictive and intelligent water management**. By combining leakage detection, leakage localization, and consumption forecasting, the system can help reduce water wastage, minimize maintenance costs, improve distribution efficiency, and support sustainable utilization of water resources. Such an approach is particularly useful for smart cities, residential communities, industrial facilities, agricultural networks, and municipal water distribution systems, where continuous monitoring and accurate prediction are essential for effective water resource management.

II. LITERATURE SURVEY

The increasing demand for water and the aging of water distribution infrastructure have encouraged researchers to develop intelligent techniques for water leakage detection, localization, and consumption forecasting. Traditional approaches based on manual inspection, acoustic equipment, pressure monitoring, and periodic meter readings are often expensive and may not detect small or developing leaks at an early stage. Recent research has therefore focused on integrating IoT sensors, machine learning, deep learning, hydraulic modeling, and data-driven forecasting to develop smart water management systems.

1. Water Leakage Detection Using Machine Learning

Machine learning has been increasingly applied to identify abnormal patterns in water distribution networks. Conventional statistical methods mainly rely on predefined thresholds for flow and pressure. However, these thresholds may produce false alarms because normal water demand can vary significantly with time and location. Machine learning models can learn normal consumption behavior from historical data and identify deviations that indicate possible leakage.

Supervised algorithms such as Decision Trees, Random Forest, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Gradient Boosting have been investigated for classifying water-network conditions. These algorithms can use parameters such as inlet flow, outlet flow, pressure, water demand, and time-based features to distinguish normal conditions from leakage events. Ensemble learning approaches are particularly useful because they can combine multiple decision boundaries and handle nonlinear relationships between sensor parameters.

2. IoT-Based Smart Water Monitoring

The development of the Internet of Things has enabled continuous monitoring of water distribution systems. IoT-based systems commonly employ flow meters, pressure sensors, smart meters, and wireless communication technologies to collect real-time information from pipelines. Sensor measurements can be transmitted to a centralized server or cloud platform for analysis. IoT-based monitoring provides several advantages over conventional manual inspection because water-network parameters can be observed continuously. When sensor readings show abnormal changes, the system can automatically generate an alert. However, IoT systems can produce large amounts of heterogeneous data, and sensor noise, missing readings, communication failures, and calibration errors can affect detection accuracy. Machine learning provides an effective mechanism for processing these large-scale sensor datasets.

3. Leakage Localization in Water Distribution Networks

Detecting a leak is only the first step in effective water-network management. Identifying the probable location of the leakage is equally important because maintenance teams need to know where to inspect or repair the pipeline. Research in leakage localization has explored hydraulic modeling, pressure-based methods, sensor networks, and machine learning. Pressure changes at different locations can provide useful information about a leak. A sudden pressure reduction may indicate leakage in a particular network section. By analyzing pressure and flow variations from multiple sensors, machine learning models can learn the relationship between sensor measurements and leak locations. Classification models can then assign a detected leakage event to a particular pipeline segment or geographical zone.

Some research has also investigated hybrid approaches in which hydraulic simulation is combined with machine learning. Hydraulic models provide knowledge about physical water-flow behavior, while machine learning learns complex patterns from simulated or real-world data. Such hybrid systems can improve localization performance but may require accurate network topology and hydraulic parameters.

4. Deep Learning for Water Leakage Detection

Deep learning techniques have also been explored for complex water-network monitoring problems. Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs) can learn nonlinear relationships from large datasets. RNN-based models, particularly Long Short-Term Memory (LSTM) networks, are useful when water-flow and consumption data have strong temporal dependencies. An LSTM model can analyze historical sensor measurements and learn how water-network behavior changes over time. Abnormal deviations between expected and actual behavior can subsequently be used to identify potential leakage. Deep learning can provide higher predictive capability for complex datasets, but it generally requires a larger amount of training data and greater computational resources than conventional machine learning algorithms.

5. Water Consumption Forecasting

Water consumption forecasting is another important research area in smart water management. Accurate prediction of future demand helps water utilities plan supply, optimize distribution, and identify abnormal consumption. Water demand is influenced by several factors, including time of day, day of the week, season, weather, temperature, rainfall, population, and consumer behavior. Traditional forecasting techniques such as moving averages, regression, and autoregressive models have been widely used. More recently, machine learning approaches such as Random Forest, Support Vector Regression, Gradient Boosting, Artificial Neural Networks, and LSTM have been employed to model nonlinear consumption patterns. LSTM-based forecasting is particularly suitable for water consumption because consumption measurements form time-series data. The model can learn relationships between previous observations and future demand. Forecasting can also contribute to leakage detection: if the actual consumption is significantly higher than the predicted consumption without an identifiable reason, the difference can be treated as a potential anomaly.

6. Smart Meter-Based Consumption Analysis

Smart water meters provide detailed consumption measurements at relatively short time intervals compared with traditional monthly meter readings. Researchers have used smart-meter data to identify consumption patterns, detect unusual behavior, and forecast demand. Consumption profiles can be analyzed at household, commercial, or district levels. Machine learning algorithms can categorize consumers according to their usage patterns and identify deviations from their normal profiles. This approach can help distinguish legitimate high consumption from potentially abnormal consumption caused by leaks. However, smart-meter-based systems can face challenges such as incomplete datasets, irregular sampling, privacy concerns, sensor failures, and variations in consumer behavior.

7. Anomaly Detection for Water Distribution Networks

Anomaly detection is closely related to leakage detection. Instead of directly training a model using only labeled leakage events, anomaly detection methods attempt to learn the normal operating behavior of a water network. Any significant deviation from the learned behavior can then be flagged as an anomaly. Both supervised and unsupervised learning approaches have been considered. Supervised learning can provide high accuracy when sufficient labeled leak data are available. However, obtaining real-world labeled leakage datasets can be difficult because utilities cannot deliberately create large numbers of real leakage events. Unsupervised techniques such as clustering, isolation-based methods, and auto encoders can therefore be useful when labeled data are limited.

III. EXISTING SYSTEM

In existing water-distribution networks, water leakage detection and consumption monitoring are primarily performed using manual inspection, periodic meter readings, threshold-based monitoring, and conventional hydraulic analysis. Water utilities and building maintenance teams generally monitor water flow and pressure at selected points and inspect pipelines when abnormal readings, visible water accumulation, reduced pressure, or increased water bills are observed. In many cases, underground pipeline leakage is identified only after significant water loss or infrastructure damage has occurred. Traditional leakage detection methods commonly use acoustic sensors, pressure gauges, flow meters, ground microphones, tracer techniques, and visual inspection. These methods can help identify leakage, but they often require specialized equipment and experienced personnel. Since pipelines may be underground or located in inaccessible areas, continuous physical inspection is difficult and costly. Small leaks may therefore remain undetected for extended periods. Some existing systems use fixed threshold-based techniques for detecting abnormal water flow or pressure. For example, if the measured flow exceeds a predefined limit or the pressure falls below a specified threshold, the system generates a leakage warning. Although such approaches are simple to implement, they may produce false alarms because water consumption and pipeline pressure naturally vary according to time, demand, weather, and operational conditions. Fixed thresholds are also less effective for detecting gradual or small leakage events.

DISADVANTAGES OF EXISTING SYSTEM

- Delayed Leakage Detection.
- Limited Leakage Localization.
- Dependence on Fixed Thresholds.

IV. PROPOSED SYSTEM

The proposed Machine Learning-Based Smart Water Leakage Detection, Localization and Consumption Forecasting System is an intelligent IoT-enabled framework designed to continuously monitor water-distribution networks, detect leakage at an early stage, identify the probable leakage zone, and forecast future water consumption. The system combines IoT sensors, ESP32-based communication, real-time data processing, machine learning algorithms, time-series analysis, and a web-based monitoring dashboard into a single integrated platform. Flow and pressure sensors are installed at selected locations in the water pipeline network. These sensors continuously measure parameters such as water flow rate, inlet and outlet flow, pipeline pressure, and consumption**.** The sensor readings are collected by an ESP32 microcontroller, which acts as the communication and data-acquisition unit. The collected data are transmitted to a centralized server or cloud-based monitoring platform for further processing and analysis. The collected sensor data are first subjected to data preprocessing to handle missing values, noise, abnormal readings, and inconsistencies. The processed time-series data are then transformed into meaningful features, including flow imbalance, pressure variation, rate of change, rolling averages, consumption lags, and temporal features**.** These features provide a more informative representation of the operating condition of the water network. For leakage detection, machine-learning algorithms such as Random Forest, Support Vector Machine (SVM), and Gradient Boosting are trained using historical normal and leakage-related data. The models classify the current water-network condition into categories such as normal operation, small leakage, medium leakage, and major leakage. Model performance can be evaluated using accuracy, precision, recall, F1-score, and confusion matrix analysis, allowing the most suitable model to be selected.

ADVANTAGES

- Early Leakage Detection.
- Real-Time Monitoring.
- Accurate Leakage Classification.
- Leakage Localization.

SYSTEM ARCHITECTURE

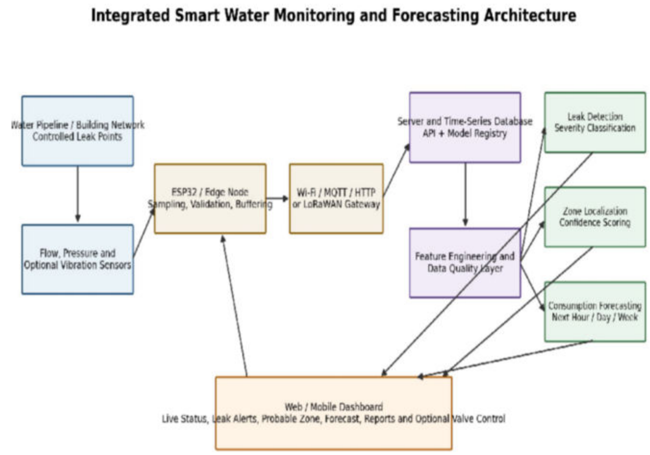

Fig 1: System Architecture

V. UML DIAGRAMS

1. ACTIVITY DIAGRAM

An Activity Diagram is a UML diagram. It represents the workflow of a system. The diagram shows the sequence of activities. These activities are performed from the start to the end of a process. It helps to understand how tasks are done. Also it shows how data moves through stages.

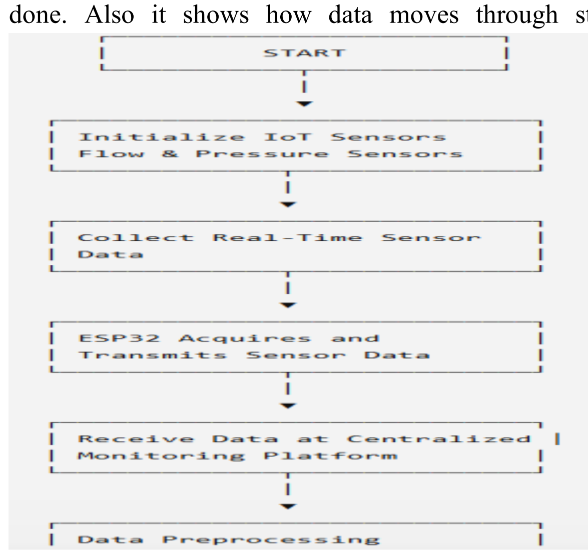

Fig 5.1 shows the Activity diagram

2. USECASE DIAGRAM:

A Use Case diagram describes the structure of a system by showing its classes, attributes, methods, and the relationships between them (like inheritance or association). It acts as the blueprint of object-oriented design and helps understand how data and functions are organized.

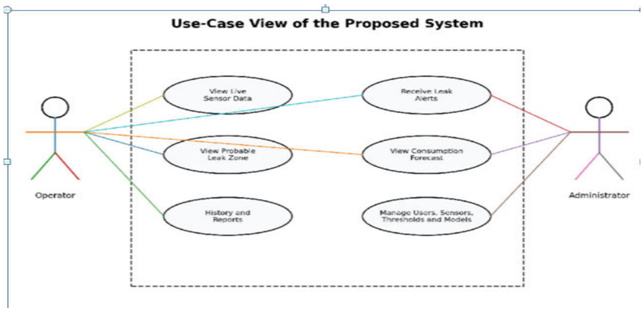

Fig 5.2 Shows the Use case Diagram

3. SEQUENCE DIAGRAM:

A sequence diagram shows how different objects communicate over time. It uses a time-ordered sequence of messages to explain interactions step by step, making it useful for understanding the flow of logic in a scenario.

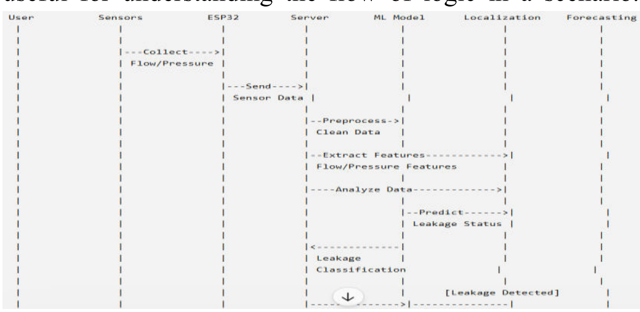

Fig 5.3 Shows the Sequence Diagram

VI. RESULTS

6.1 Output Screens

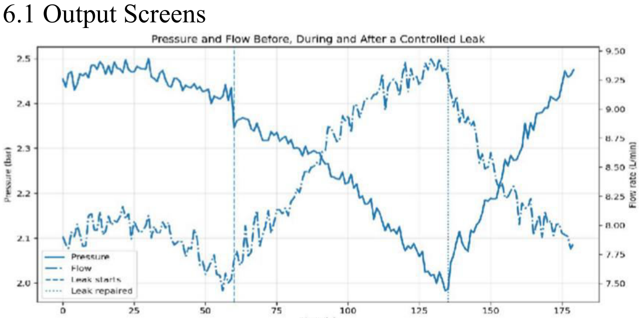

Fig 6.1 Pressure and flow signals before and after a controlled leak

Pressure and flow signals before and after a controlled leak.

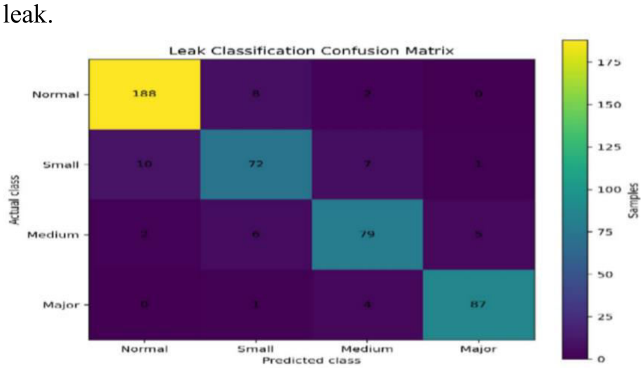

Fig 6.2 Leak Classification Confusion Matrix

In above screen shows the Leak Classification Confusion Matrix.

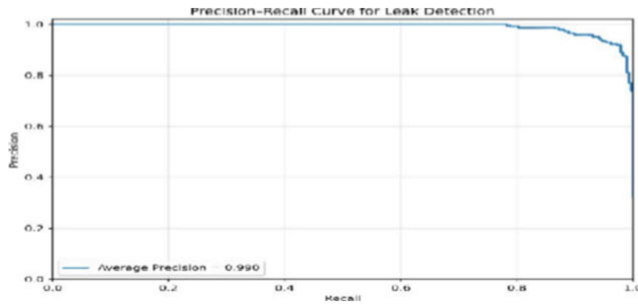

Fig 6.3 precision – recall curve for leak Detection

In above screen shows the precision – recall curve for leak Detection

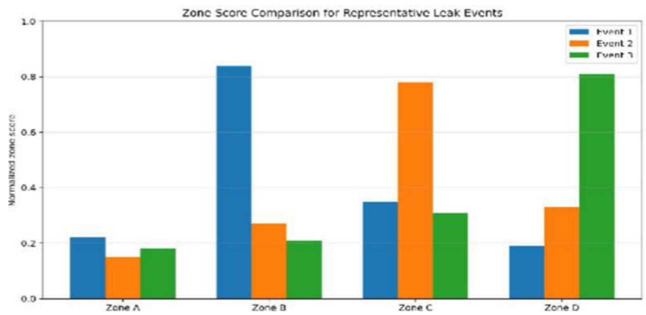

Fig 6.4 Zone-Score Comparison for Representative Leak Events

In above screen shows the Zone-Score Comparison for Representative Leak Events.

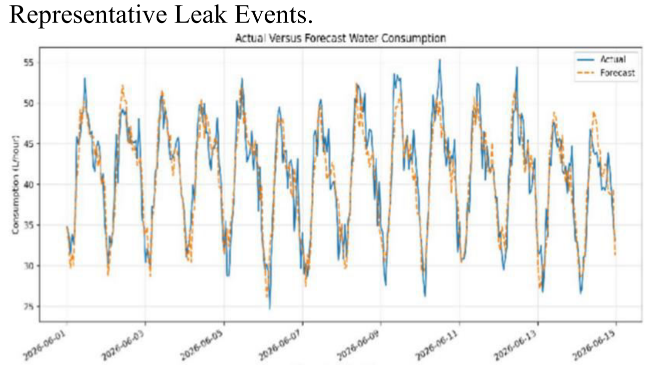

Fig 6.5 Actual versus Forecast Water Consumption

In above screen shows the Actual versus Forecast Water Consumption

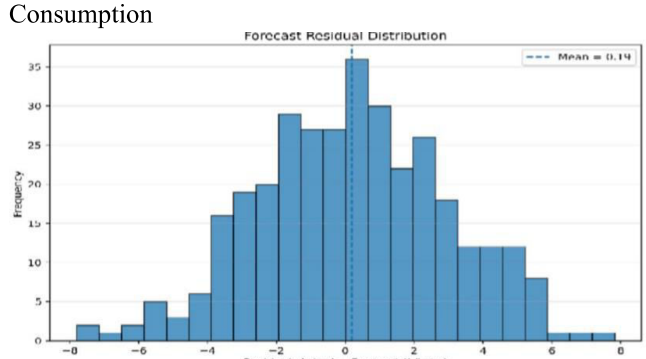

Fig 6.6 Forecast Residual Distribution

In the above screen shows the Forecast Residual Distribution

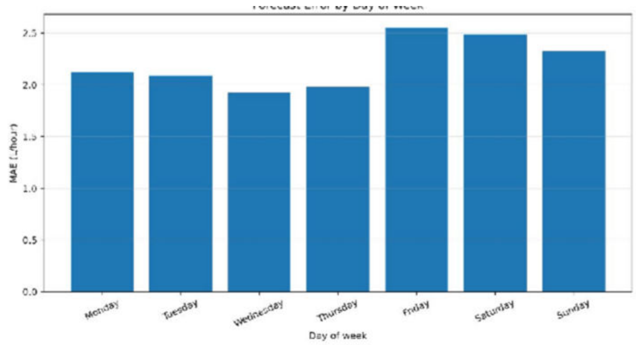

Fig 6.7 Forecast Error By Day of Week

In above screen shows the Forecast Error by Day of Week.

VII. CONCLUSION

This project proposes a practical and academically complete architecture for machine-learning-based smart water leakage detection, zone localization and consumption forecasting. Pressure and flow sensors provide continuous information about pipeline behavior, while IoT nodes deliver validated time-series data to a server. A hybrid leak engine combines hydraulic residuals with supervised classification, and a zone-scoring method reduces the inspection area for maintenance. A separate forecasting engine learns lagged, seasonal and contextual patterns to estimate future aggregate demand. By sharing data, storage, APIs and visualization, the three functions form a unified smart-water platform rather than isolated demonstrations. The project is intentionally designed to be scalable but honest about its limits. It does not claim exact leak coordinates or field-level accuracy without evidence. Final submission should replace all expected values with measurements from the completed prototype, state the dataset size and split method, compare against baselines and clearly distinguish simulated, laboratory and field results.

VIII. REFERENCES

[1] M. R. Islam, S. Azam, B. Shanmugam, and D. Mathur, "A Review on Current Technologies and Future Direction of Water Leakage Detection in Water Distribution Network," IEEE Access, vol. 10, pp. 107177-107201, 2022, doi: 10.1109/ACCESS.2022.3212769.

[2]R.J.L. Taloma, F.Cuomo, D.Comminiello, and P.Pisani, " Machine Learning for Smart Water Distribution Systems: Exploring Applications, Challenges and Future Perspectives," Artificial Intelligence Review, vol. 58, art. 120, 2025, doi: 10.1007/s10462-024-11093-7.

[3] T. Tanchangya, A. Raihan, J. Rahman, and M. Ridwan, "A Review of Deep Learning Applications for Sustainable Water Resource Management," Global Sustainability Research, vol. 3, no. 4, 2024, doi:10.56556/gssr.v3i4.1043.

[4] G. Rajan and S. Li, "A Systematic Literature Review on Flow Data-Based Techniques for Automated Leak Management in Water Distribution Systems," Smart Cities, vol. 8, art. 78, 2025, doi:10.3390/smartcities8030078.

[5] L.Romero-Ben, F.Alves, D.Blesa, G. Cembrano, V.Puig, and E.Duviella, " Leak Detection and Localization in Water Distribution Networks: Review and Perspective, " Annual Reviews in Control,vol.55,pp.392-419,2023.

[6] H. Gemeinhardt et al., "Machine-Learning-Assisted Leak Detection Using Distributed Acoustic Sensing," IEEE Transactions on Instrumentation and Measurement, vol. 72, 2023.

[7] C. U. Onukwube, D. O. Aikhuele, and S. Sorooshian, "Development of a Fault Detection and Localization Model for a Water Distribution Network," Applied Sciences, vol. 14, no. 4, art. 1620, 2024, doi:10.3390/app14041620.

[8]N.M. Mahdietal. , " Leak Detection and Localization in Water Distribution Systems Using Advanced Feature Analysis and an Artificial Neural Network, " Desalination and Water Treatment, vol.320,art.100685,2024.

[9] M. Adraoui et al., "A Two-Phase Approach for Leak Detection and Localization in Water Distribution Systems Using Wavelet Decomposition and Machine Learning," Computers & Industrial Engineering, art. 110534, 2024.

[10] N. Ullah, M. F. Siddique, S. Ullah, Z. Ahmad, and J.-M. Kim, "Pipeline Leak Detection System for a Smart City: Leveraging Acoustic Emission Sensing and Sequential Deep Learning , " SmartCities,vol.7,no.4,pp. 2318-2338, 2024.

[11] F. Saleem, Z. Ahmad, and J.-M. Kim, "Real-Time Pipeline Leak Detection: A Hybrid Deep Learning Approach Using Acoustic Emission Signals," Applied Sciences, vol. 15, no. 1, art. 185, 2025.

[12]S.G. Vrachimisetal., " Battle of the Leakage Detection and Isolation Methods, " Journal of Water Resources Planning and Management, vol. 148, no. 12, art. 04022068, 2022.

[13] S.Wu, P.Zhao, M.Bai, J.Wang, and Y.Lan, "A Fine Grained Forecasting Method for Minute-Level Urban Water Demand," in Proc. 40th Chinese Control Conference, 2021, pp. 8298-8303.

[14] P.Hu, J.Tong, J.Wang, Y.Yang, and L.Oliveira Turci, "A Hybrid Model Based on CNN and Bi-LSTM for Urban Water Demand Prediction," in Proc. IEEE Congress on Evolutionary Computation, 2019, pp. 1088-1094.

[15] D.Kofinas, E.Papageorgiou, C.Laspidou, N.Mellios, and K.Kokkinos, " Daily Multivariate Forecasting of Water Demand in a Touristic Island with Artificial Neural Network and Adaptive Neuro-Fuzzy Inference System," in Proc. CySWater, 2016, pp. 37-42.
[16] A. Zanfei, B. M. Brentan, A. Menapace, M. Righetti, and M. Herrera, "Graph Convolutional Recurrent Neural Networks for Water Demand Forecasting," Water Resources Research, vol. 58, no. 7, 2022.
[17] D. Li and Y. Fu, "Deep Learning Model-Based Demand Forecasting for Secondary Water Supply in Residential Communities: A Case Study of Shanghai City, China," IEEE Access, 2023.