IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 4, June 2025

Impact Factor: 7.67

Water Consumption Prediction Using Machine Learning Algorithm

Prof. S. S. Bhosale[1], **Gadekar Pooja**[2], **Pandagale Prasad**[3], **Badakh Vijeta**[4], **Chaudhary Chaitanya**[5]

Faculty, Department of Information Technology [1]

Students, Department of Information Technology [2,3,4,5]

Pravara Rural Engineering College,Loni,India

Abstract: *Water is a crucial resource, and regulating its consumption is critical to long-term sustainability .Understanding and anticipating water usage trends has become increasingly important for resource management, conservation, and strategic planning as populations and cities develop. This research analyses and forecasts water use using a predictive model based on the Support Vector Machine (SVM) algorithm and a textual collection of historical data. This technology may predict future demand by identifying water usage patterns, allowing for more efficient water distribution and preventing short ages To provide accurate forecasts about water usage, the system uses important procedures such as feature extraction, segmentation, and classification.*

Keywords: Water Prediction, Machine Learning, Random Forest, Linear Regression

I. INTRODUCTION

The introduction of the water consumption prediction system focuses on the critical need for effective water resource management in response to increasing demand. As populations grow and urbanization expands, the strain on water resources has intensified, making it essential to understand and anticipate consumption patterns. Traditional methods of water management are often reactive, relying on historical usage patterns without the predictive insights necessary for proactive planning. This project addresses these challenges by developing a predictive model based on Support Vector Machine (SVM) algorithms, capable of Analyzing historical water usage data to forecast future consumption. By accurately classifying and predicting usage patterns, this system can assist municipalities, industries, and residential areas in making informed decisions regarding water distribution, conservation, and infrastructure planning. The water consumption prediction system thus aims to transform resource management from a reactive to a proactive approach, ultimately supporting sustainability and efficient resource allocation

II. LITERATURE REVIEW AND OBJECTIVE

A. LITERATURE REVIEW

The issue of domestic water consumption has been explored in various studies, focusing on different aspects and technologies that aim to monitor, manage, and reduce water usage in households and buildings. One such study, Defining Domestic Water Consumption Based on Personal Water Use Activities by Melissa Lauren Crouch, Heinz Erasmus Jacobs, and Vanessa L. Speight (2023), shifts the focus from traditional household consumption metrics to the quantification of individual water use. This research highlights the importance of understanding personal water use activities, regardless of the water source or location, which is crucial for accurately assessing the absolute basic consumption (ABC) of an individual.

In a similar vein, the Water Consumption Monitoring System developed by Tharun Kumar G. and Dr. R. Puviarasi (2021) aims to address the need for monitoring and managing domestic water consumption through advanced systems. This technology provides real-time tracking of water usage in households, promoting water conservation and efficient management. However, as the study by Lili Yang, Shuang-Hua Yang, Ewa Magiera, Wojciech Froelich, Tomasz Jach, and Chrysi Laspidoud (2021) points out, the trend of increasing individual households has led to less efficient water usage, necessitating innovative solutions like Internet of Things (IoT) technologies for behavior intervention. The

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, June 2025
Impact Factor: 7.67

application of IoT in monitoring domestic water consumption not only provides real-time data but also helps implement behavior modification strategies to promote water conservation.

A more ambitious approach is explored by Cheng-li Cheng, Sadahico Kawamura, and Wei-Che Chang (2021) in their study on Research on a Zero Water Consumption Operation Model and Feasibility for Office Buildings. With Taiwan's government targeting net-zero emissions by 2050, this study investigates the feasibility of adopting zero water consumption designs for office buildings, aligning with global sustainability goals. The transition toward zero water consumption and energy-efficient buildings plays a vital role in achieving these environmental objectives.

Finally, the Urban Water Consumption: A Systematic Literature Review by Talita Flores Dias and Enedir Ghisi (2021) reviews the main factors influencing water consumption in urban settings. This comprehensive review helps researchers and urban managers understand the patterns of water usage across different regions, highlighting the diverse factors that contribute to water consumption trends and providing valuable insights for future water management strategies.

Together, these studies reflect the growing need for advanced monitoring systems, individual behavior interventions, and sustainable building designs to address the pressing issue of domestic water consumption. Each study provides a unique contribution, whether through the application of IoT technologies, a shift in focus from household to individual consumption, or the exploration of zero water consumption models, all aiming toward a more efficient and sustainable use of water resources.

III. PROPOSED SYSTEM

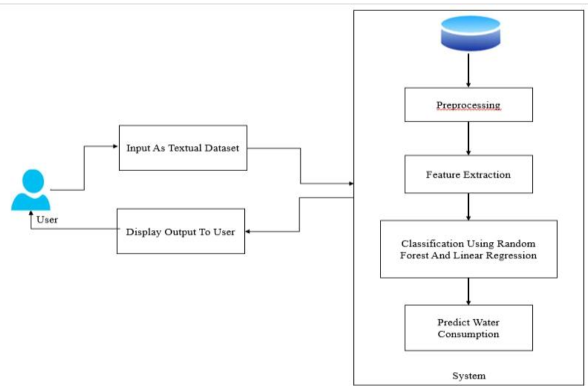


FIG 1: OVERVIEW SYSTEM ARCHITECTURE

1. Data Collection & Preprocessing: Collect historical water usage data, clean it, and prepare it for analysis.
2. Feature Extraction: Identify relevant features that influence water consumption patterns.
3. Data Segmentation: Segment data by time periods (e.g. monthly , yearly) or other attributes to refine analysis.
4. Training Model: Use Linear Regression and Random Forest algorithm to train the model and analyze water usage patterns.
5. Prediction Module: Apply the trained model to forecast future water consumption based on learned patterns.
6. Classifiers: To evaluate the effectiveness of water consumption pattern, both Linear Regression and Random Forest approaches are considered.

Linear Regression Algorithm : Linear Regression is a statistical model that predicts an outcome (y) based on one or more input variables(x) using a straight line equation.

1. Simple linear regression

$y=\beta 0+\beta 1X$

2. Multiple linear regression $y=\beta 0+\beta 1X1+\beta 2X2+\ldots\ldots\ldots\beta nXn$

How it Works : Identifies a linear relationship between input and output. Fits a straight line my minimizing the error Predict feature values using learned equation

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, June 2025
Impact Factor: 7.67

.
Random Forest Algorithm : Random Forest is a machine learning algorithm that makes predictions by combining multiple decision trees. Instead of relying on a single tree, it builds many trees and takes the average result for better accuracy.
How it works : Split data into multiple decision trees Each tree predicts an output The model average for the final prediction.

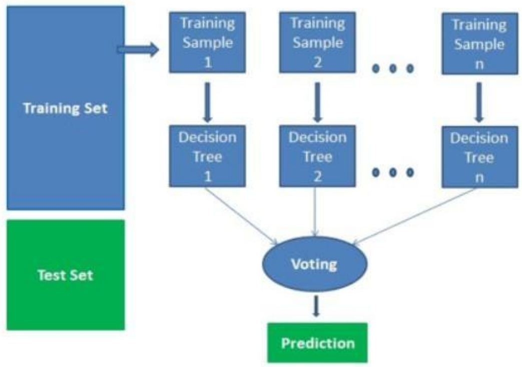


Fig 2. Random Forest

IV. ANALYSIS OF THE WATER CONSUMPTION PATTERN CLASSIFICATION

4.1. Pre-processing :

Pre-processing for water consumption prediction using machine learning involves several key steps aimed at preparing data for accurate modelling . First, relevant data is collected, including historical consumption, weather conditions, and temporal factors. The data is then cleaned by handling missing values, detecting and removing outliers, and ensuring consistency. Feature engineering follows, where new variables are created, such as time-based features and weather-related ones, to capture patterns and dependencies in water usage. Data is normalized or standardized, and categorical variables are encoded for machine learning algorithms. Feature selection is performed to reduce redundancy and focus on the most important variables. The dataset is split into training, validation, and test sets, and techniques like cross-validation are used to assess model performance. If data imbalance exists, strategies like oversampling or under sampling are applied. Effective pre-processing ensures the model is trained on high-quality data, improving its accuracy in predicting water consumption patterns.

4.2. Feature Extraction :

Feature extraction for water consumption prediction involves transforming raw data into relevant, informative features that capture the underlying patterns in water usage. Key factors influencing consumption include time-based patterns, such as daily, weekly, and seasonal cycles, which reflect regular fluctuations in usage. Weather conditions, like temperature, rainfall, and humidity, also play a significant role in determining water demand, particularly for irrigation or cooling. Temporal dependencies, such as past consumption, help models predict future usage based on historical data. Additionally, aggregating consumption data over time, identifying special events or holidays, and considering demographic and socio-economic factors—such as household size or population density—further enhance the model's predictive power. Geographic features, like water availability and region-specific characteristics, along with external factors like water pricing and regulations, also influence consumption patterns. Effective feature extraction enables machine learning models to understand complex consumption behaviors, improving prediction accuracy and reliability.

4.3. Data Segmentation :

Data segmentation is a crucial process in improving the accuracy of machine learning models for predicting water consumption patterns. By dividing the data into meaningful subsets based on temporal, geographic, demographic,

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, June 2025
Impact Factor: 7.67

behavioral, and external factors, machine learning models can better capture the diverse influences on water usage. Temporal patterns, such as daily, weekly, and seasonal cycles, help identify predictable fluctuations, while geographic segmentation accounts for regional differences in consumption. Demographic factors allow the model to distinguish between residential, commercial, and industrial users, while behavioral segmentation uncovers unique consumption trends. External factors like weather and regulatory measures further refine predictions. This segmentation enables models to focus on relevant patterns within specific data subsets, leading to more accurate and actionable predictions for water management and consumption forecasting.

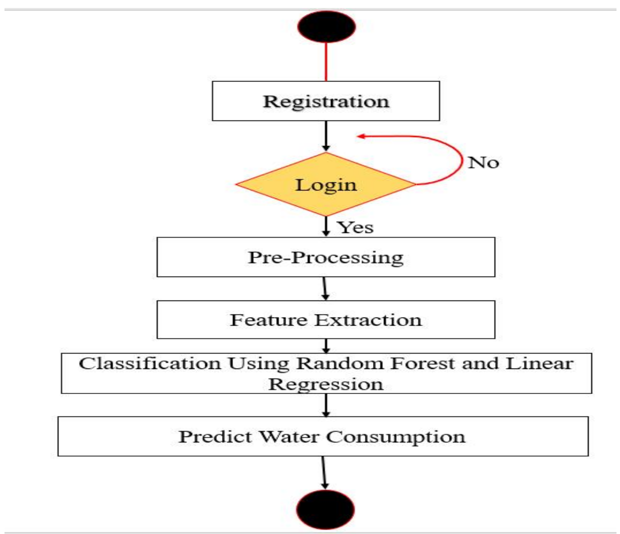
4.4 Training Model :

This research focuses on predicting water consumption patterns using Linear Regression and Random Forest models. Data is collected from sources like smart meters, weather reports, and demographic statistics. After preprocessing and feature engineering, Linear Regression helps analyze the impact of individual factors but assumes a linear relationship. In contrast, Random Forest captures complex, non-linear patterns by combining multiple decision trees for better accuracy. Model performance is evaluated using MAE, RMSE, and R^2 scores. The trained model can be deployed for real-time monitoring, aiding in sustainable water management and efficient resource distribution.

4.5. Implementation info :

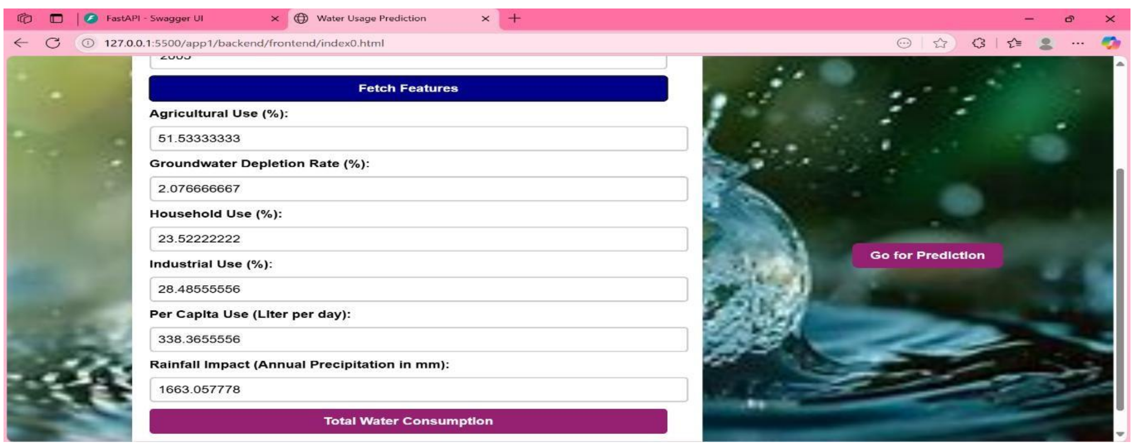
1. Hardware Requirements :Processor: Intel i5 or above RAM: 8 GB minimum (16 GB recommended for large datasets)Storage: 250 GB for data storage and processing
2. Software Requirements Operating System: Windows, macOS, or Linux, Programming Language - Python 3.x , Libraries: Machine Learning - scikit-learn (for SVM) Data Processing - pandas, numpy ,Data Visualization - matplotlib, seaborn, Database: SQLite for data storage , IDE: Jupyter Notebook, VS Code, or PyCharm for development
3. Additional Requirements : Internet: Required for accessing online datasets or libraries , Data Sources: Historical water consumption datasets for model training and testing.

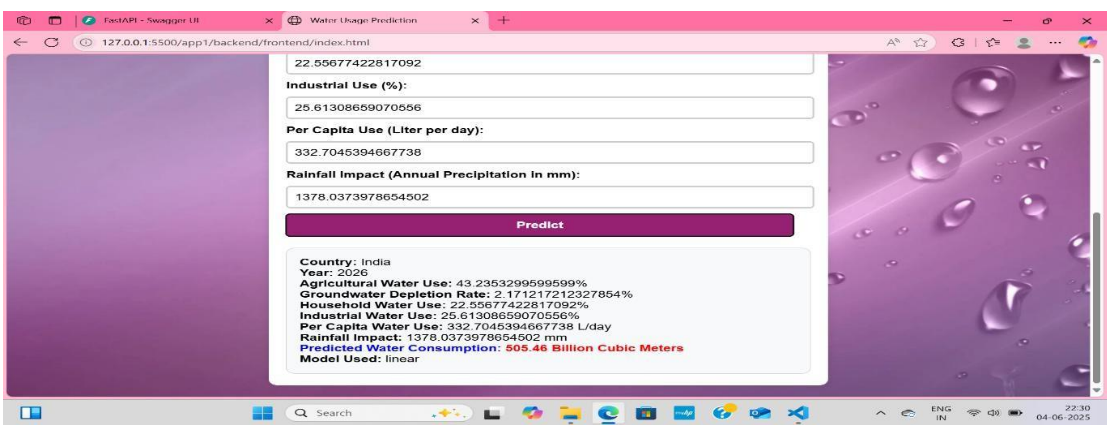
V. FLOWCHART AND ALGORITHM



IJARSCT
ISSN: 2581-9429
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 5, Issue 4, June 2025
Impact Factor: 7.67

VI. RESULT





DOI: 10.48175/IJARSCT-27636

VII. CONCLUSION

The Our project utilizes machine learning algorithms to predict water consumption, enabling efficient resource management and anomaly detection. The model improves forecasting accuracy, supporting sustainable water usage. Future enhancements may include real-time data integration for better decision-making.

REFERENCES

[1] Defining domestic water consumption based on personal water use activities Melissa Lauren Croucha, Heinz Erasmus Jacobsa and Vanessa L. Speight b,* a Department of Civil Engineering, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa b Department of Civil and Structural Engineering, University of Sheffield, Sir Frederick Mappin Building, Mappin Street, Sheffield S1 3JD, UK *Corresponding author. E- mail: v.speight@sheffield.ac.uk.

[2] “WATER CONSUMPTION MONITORING SYSTEM” Tharun kumar.G[1], Dr.R.Puviarasi[2] Electronics and communication engineering, Saveetha school of engineering, SIMATS, Chennai-602105 tharunkumar1307@gmail.com,puviarasi@saveetha.com International Conference on Recent Trends in Computing, Communication and Networking Technologies (ICRTCCNT’19) Oct 18-19, 2019, Kings Engineering College, Chennai, TamilNadu, India.

[3] International Conference on Advances in Information Technology, IAIT2016, 19-22 December 2016, Macau, China Domestic water consumption monitoring and behaviour intervention by employing the internet of things technologies Lili Yanga*, Shuang-Hua Yangb, Ewa Magierac, Wojciech Froelichc, Tomasz Jachc, Chrysi Laspidoud a School of Business and Economics, Loughborough University, Loughborough, LE11 3TU, UK b School of Business and Economics, Loughborough University, Loughborough, LE11 3TU, UK c Institute of Computer Science, University of Silesia, Sosnowiec, 41-200, Poland d Centre For Research and Technology Hellas, Greece