Journal of Engineering Issues and Solutions

Predictive Water Usage Optimization with IoT and Regression Models

Yuyutshu Banjara[1*], Dhiraj KC[2,]

[1] *Nepal Open University, Nepal*
[2] *Purwanchal Campus, IOE, Tribhuban University*
**Corresponding email: uutshu@hotmail.com*
Submitted: April 13, 2025; Revised: June 15, 2025; Accepted: June 21, 2025

https://doi.org/10.3126/joeis.v4i1.81612

Abstract

Water scarcity and inefficient water distribution systems continue to pose significant global challenges, particularly in regions with limited infrastructure. This study proposes an IoT-enabled water management system that integrates regression-based predictive modeling to optimize consumption and reduce wastage. The system employs ultrasonic sensors to continuously monitor water tank levels and gathers real-time usage data. Using this data, multiple machine learning algorithms including Linear Regression, Polynomial Regression, and Random Forest Regression were evaluated to forecast short-term water demand. Among them, the Polynomial Regression model of degree 4 achieved the highest accuracy with an R^2 score of 0.727.

One of the system's key innovations lies in its ability to locally update the predictive model on the NodeMCU microcontroller. A weighted update mechanism is employed, where the stored model retains a cumulative weight representing the past 30 days, and each new data point is added with a weight of 1. This approach ensures the model remains responsive to recent consumption patterns while maintaining historical trends—allowing lightweight, on-device forecasting without needing constant cloud retraining.

The predicted consumption values feed into a smart refilling mechanism, automating water distribution based on anticipated demand. This scalable and cost-effective solution supports sustainable water management by embedding real-time intelligence directly into low-power devices.

Keywords: *IoT, Polynomial Regression, Demand Forecasting, Water Management, Edge Computing*

1. Introduction

Water shortage and inefficient water use at home are serious global problems, made worse by growing city populations and climate change (Abdel Nasser, Rashad, & Hussein, 2020). Conventional manual water control methods often cause water wastage or shortages. To fix this, smart and automatic systems are being developed that can measure water levels and predict future use.

IoT-based systems use ultrasonic sensors to continuously measure water levels without touching the water directly (Huque, Jafor, Pushon, & Ekah, 2023; Sundaralingapandi, Siva, & Hariprasath, 2023). These systems can automatically turn pumps on or off based on how much water is needed, helping to save water and avoid shortages.

However, many current systems depend on cloud servers to process data, which can cause delays, use more power, and create security risks (Abdel Nasser et al., 2020; Xu et al., 2020). Recently, running smart algorithms directly on small devices (called edge intelligence) is becoming popular because it is faster and more efficient (Xu et al., 2020).

This study depicts a fully automatic IoT water management system that uses a degree-4 polynomial regression model running on a NodeMCU microcontroller. This model updates its predictions every day using data from the last 30 days, allowing it to adapt to changes in water use over time without needing to connect to the cloud. It is cost-effective, easy to scale, and can be expanded in the future to detect leaks or connect with smart home systems.

2. Methodology

This section describes the methodology adopted to design and implement an IoT-based water management system with regression-driven prediction and embedded self-updating capability. The approach is aligned with the project objectives: (1) to automate tank refilling, and (2) to forecast water consumption. The methodology integrates sensor-based data acquisition, model training and evaluation, and embedded AI-based automation.

2.1 System Overview

The system architecture is designed around an IoT framework using NodeMCU, ultrasonic sensors, and relay modules. During the development phase, Python-based backend services handled sensor data logging, visualization, and model training. Upon deployment, all computations were shifted to the edge device (NodeMCU), which now hosts a lightweight, self-updating polynomial regression model.

2.2 Hardware Setup

The core hardware components used are:

- NodeMCU (ESP32): Responsible for sensor interfacing, data collection, prediction, and pump control.
- Ultrasonic Sensor: Measures water level in the tank.
- Relay Module: Switches the pump on/off based on model output.
- Water Pump: Circulates water during refilling.

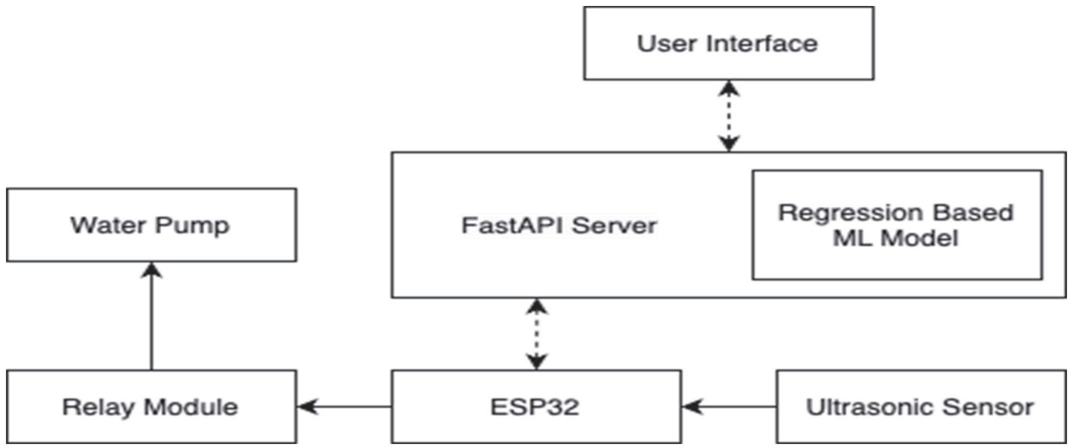

Figure 1: System during Development Phase

The sensor calculates the distance between the water and the sensor to determine the water level based on the time taken for sound waves to reflect. The level is then converted into a percentage to enable normalized data input for the prediction model.

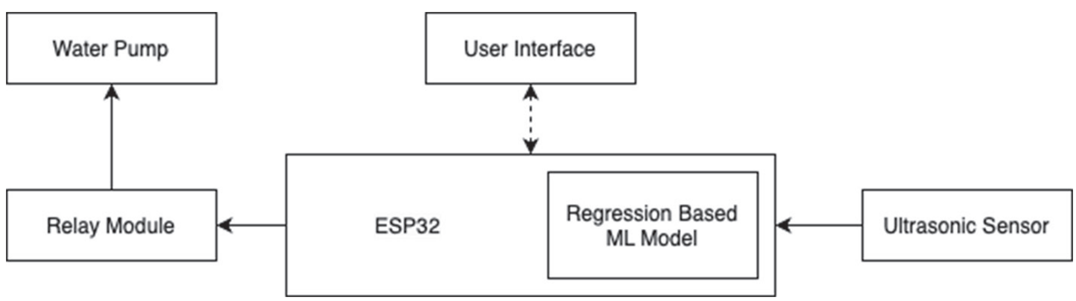

Figure 2: System during Deployment Phase

2.3 Data Collection

Sensor data (water levels) were logged every 1 hour for a period of several weeks. Daily consumption values were derived by analyzing tank depletion over time. Outliers and anomalies were filtered to preserve data quality.

2.4 Model Training and Evaluation

Three models were tested:

- Linear Regression
- Polynomial Regression (degree = 4)
- Random Forest

Polynomial Regression (degree 4) demonstrated best performance across the metrics ($R^2 = 0.727$)

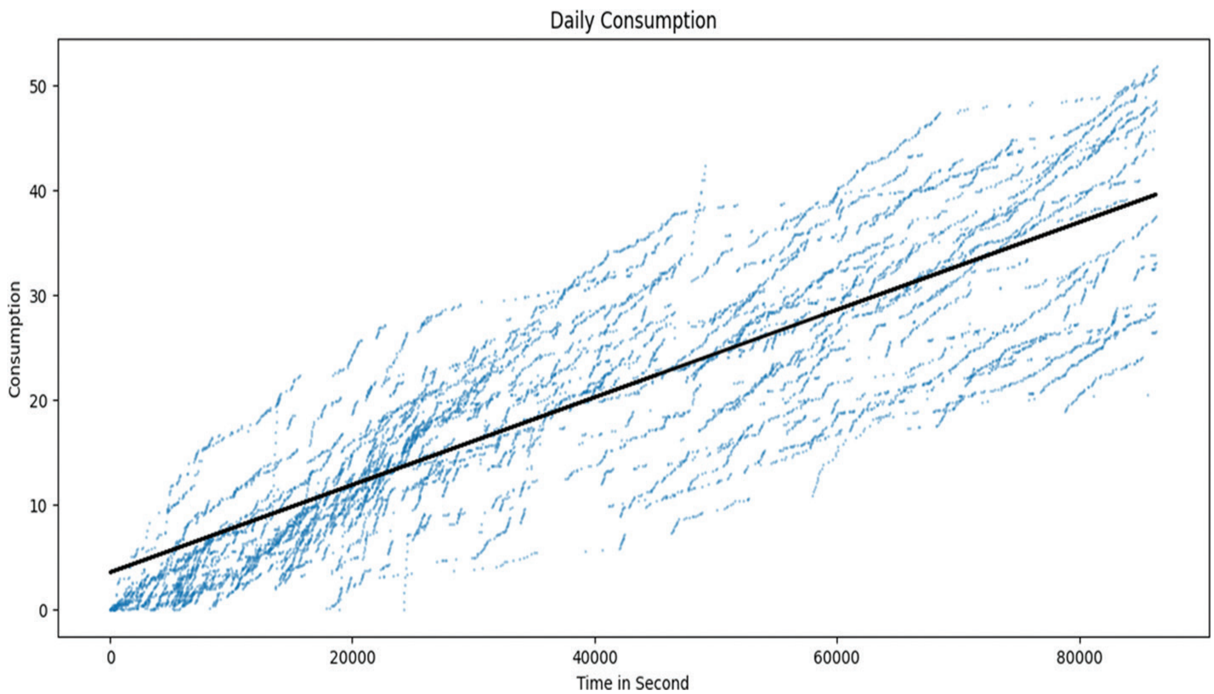

Figure 3: Linear Regression

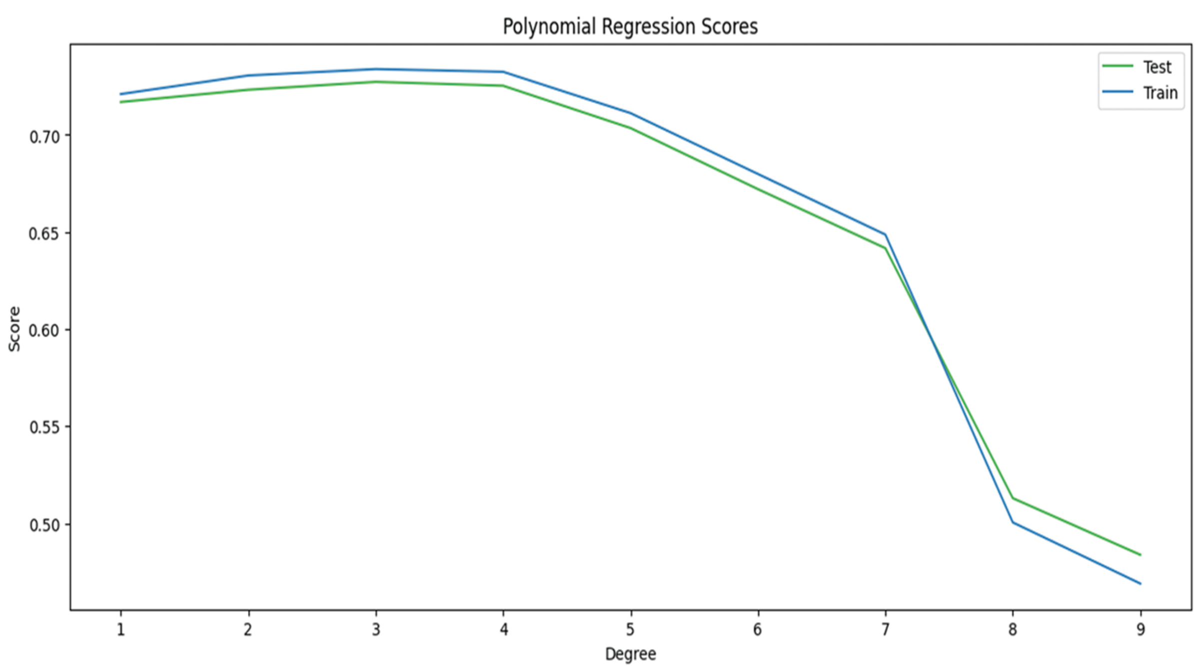

Figure 4: Polynomial Regression Scores

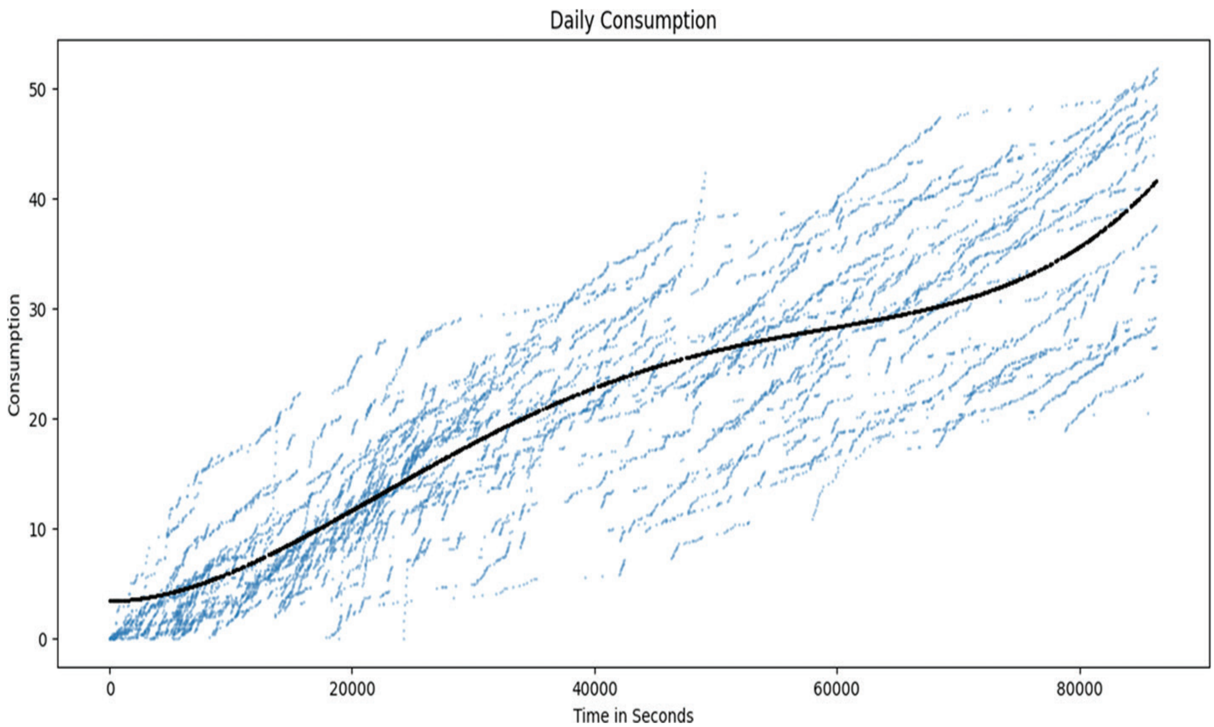

Figure 5: Polynomial Regression (degree 4)

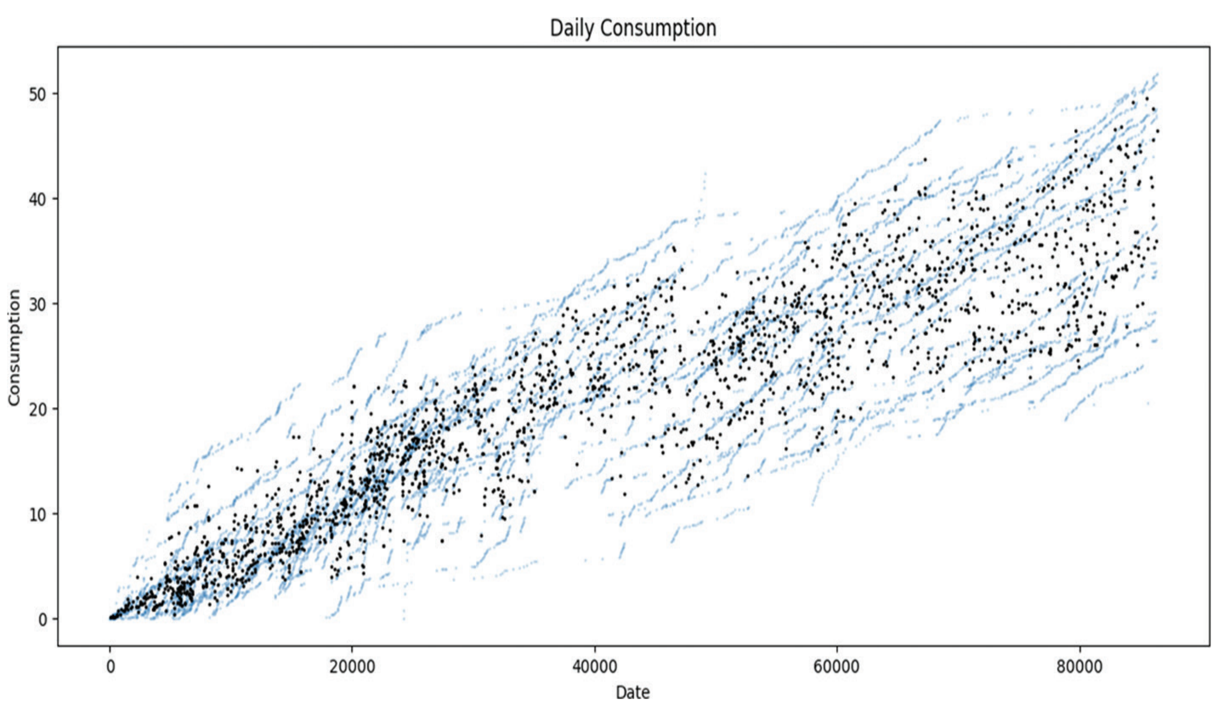

Figure 6: Random Forest

2.5 Embedded Self-Updating Model on NodeMCU

To eliminate server dependency post-deployment, the Polynomial Regression model was embedded into the NodeMCU firmware. The embedded model employs a weighted update mechanism:

- The model uses the last 30 days of data, each day assigned a weight of 1.
- The newest data point is added daily with weight 1, replacing the oldest.
- This forms a sliding-window learning algorithm with adaptive coefficients using recursive polynomial fitting.

Poly Regression Models,

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4$$

$$y_{updated} = \frac{y_{prev} \times 30 + y_{new}}{31}$$

This allows the model to adapt to seasonal and behavioral changes in water consumption patterns without external computation.

2.6 Refilling Mechanism and Logic

At user defined time, the system:

1. Predicts the next day's water demand using the polynomial model.
2. Adds a 5% safety buffer.
3. Compares it to current tank level.
4. If the water level is insufficient, it activates the pump to refill.

This allows predictive automation that ensures water availability while avoiding unnecessary pump operation.

2.7 User Interface

A SwiftUI-based frontend was created during the initial development phase, which fetched real-time data from a Python server for monitoring and visualization. However, in the deployed version, the system operates autonomously and headlessly to minimize power and memory consumption. Instead of relying on external servers or apps, the NodeMCU then functions as a lightweight HTTP server, allowing users to locally access, monitor, and edit configuration data through a simple web interface. This transition eliminated the need for continuous external connectivity, enabling an efficient, self-contained system for water management.

3. Results and Discussion

3.1 Forecasting Water Consumption Using Machine Learning

3.1.1 Regression Model Evaluation

Three machine learning models Linear Regression, Polynomial Regression (degree 4), and Random Forest Regression were trained and evaluated to forecast daily water demand based on IoT-collected data. Among the three models, Polynomial Regression of degree 4 demonstrated the best overall performance with an R^2 value of 0.727, indicating a strong fit between predicted and actual values.

Table 1: Scores with Linear, Polynomial, and Random Forest Data

Metric	Linear	Polynomial (n=4)	Random Forest
MAE	5.171	4.972	5.832
MSE	41.693	40.164	57.837
RMSE	6.457	6.337	7.605
R^2	0.716	0.727	0.607

These results showed that polynomial regression outperforms other models across all key performance indicators. It aligns with existing literature, where non-linear models were shown to better capture seasonal and behavioral trends in water consumption (Nakib et al., 2024; Arefi et al., 2021).

3.2 Real-Time Monitoring and Data Acquisition

3.2.1 Sensor Performance and Data Collection

The system relied on ultrasonic sensors interfaced with NodeMCU to gather water-level data at regular intervals. With a measurement frequency of 24 data points per day, the sensors provided real-time data which was used to calculate the water volume percentage using distance-based formulas. The sensor architecture is similar to the system proposed by Arsene et al. (2022) and Wu et al. (2023), where ultrasonic sensing and IoT integration provided robust infrastructure for water-level monitoring.

3.2.1 Self-Updating Model Algorithm

One of the key innovations in the deployed system is the self-updating mechanism of the polynomial regression model. Once deployed, the model no longer relies on an external server or re-training using Python. Instead, it is updated autonomously within NodeMCU using a weighted formula:

Updated Forecast = (Previous Model Output × 30 + New Days Prediction) ÷ 31

This formula allowed the embedded system to adapt to recent usage patterns while still retaining historical knowledge. The use of weighted updating for regression aligns with adaptive modeling principles suggested in Montgomery et al. (2021), and has been particularly suitable for low-power devices with limited processing capabilities.

3.3 Automation and System Response

3.2.1 Smart Refilling System

The refilling decision is based on daily forecasts generated by the model. At midnight (server time), the system checks the current tank level. If the level gets lower than the predicted need for the upcoming day, the pump gets automatically triggered to refill the tank to 100%.

This mechanism minimizes wastage and ensures sufficient water availability, contributing directly to more sustainable domestic resource use, in line with strategies discussed by Zhen et al. (2022) and Chopra & Meindl (2019).

3.2.1 System Optimization and Power Efficiency

To reduce power consumption, the system runs in a headless configuration post-deployment. Real-time monitoring through external applications is disabled. Instead, the NodeMCU acts as a lightweight HTTP server, allowing users to access and adjust parameters directly without requiring an always-on Python backend. This approach significantly reduces energy usage while maintaining system autonomy.

4. Conclusions

The system has been successfully developed an IoT-based water management system that integrated real-time sensing, machine learning, and embedded automation to address inefficiencies in household water usage. Using ultrasonic sensors and a NodeMCU microcontroller, the system monitored water levels and forecasted daily consumption through a self-updating polynomial regression model of degree 4, which achieved the highest accuracy ($R^2 = 0.727$) among the models tested. The model updates autonomously using a lightweight weighted algorithm that blends 30 days of prior model data with newly collected values, eliminating the need for external servers after deployment. Automated refilling based on demand predictions reduced wastage and ensured consistent water availability. The system was ultimately deployed in a headless configuration with an integrated HTTP server on the NodeMCU, enabling local access while minimizing power and memory consumption, thus achieving the projects goals of automation, forecasting, and sustainable water resource management.

References

Abdel Nasser, A., Rashad, M. Z., & Hussein, S. E. (2020). A two-layer water demand prediction system in urban areas based on micro-services and LSTM neural networks. IEEE Access, 8, 147647–147661. https://doi.org/10.1109/ACCESS.2020.3015655

Huque, M. T., Jafor, J., Pushon, S. R. R., & Ekah, F. Z. (2023). Internet of Things (IoT) based smart water tank level monitoring and motor pump control system for preventing water waste. https://www.researchgate.net/publication/371634266_Internet_of_Things_IoT_based_Smart_Water_Tank_Level_Monitoring_and_Motor_Pump_Control_System_for_Prevent_Water_Waste

Sundaralingapandi, V., Siva, R. V., & Hariprasath, R. (2023). IoT based smart water monitoring system using ultrasonic sensor. AIP Conference Proceedings, 2608, 090009. https://ui.adsabs.harvard.edu/abs/2023AIPC.2608b0009S/abstract

Xu, D., Li, T., Li, Y., Su, X., Tarkoma, S., & Crowcroft, J. (2020). Edge intelligence: Architectures, challenges, and applications. arXiv. https://doi.org/10.48550/arXiv.2003.12172