# Smart Water management in Irrigation

Subhra Jyoti, Anuleho Biswas

National Institute of Science Education and Research

**Abstract.** his study presents an innovative approach to smart water management by combining two powerful machine learning techniques - Long Short-Term Memory (LSTM) and K-Nearest Neighbors (KNN).
Our methodology involved experiments with the model plant Arabidopsis thaliana in both controlled and real-world environments. In the controlled setting, we trained the LSTM algorithm to predict soil moisture levels based on factors like temperature and humidity. We then employed KNN to classify soil conditions and refine the LSTM's predictions.
When deployed in an uncontrolled outdoor environment with varying conditions, this LSTM-KNN combination demonstrated improved accuracy in forecasting soil moisture and classifying soil states compared to traditional methods.
The results showcase the potential of integrating advanced AI techniques for precision irrigation and water conservation in agriculture. Our approach can promote sustainable farming while boosting yields by accurately predicting crop water needs and optimizing irrigation schedules.

**Keyword:** Smart irrigation, soil moisture prediction, LSTM, KNN, machine learning, precision agriculture, water conservation

## 1 Introduction

The world is facing an increasing demand for water due to population growth, urbanization, and climate change. The agriculture sector, being the largest consumer of global freshwater resources, has a significant role to play in water conservation. **Smart Water Management for Irrigation using Machine Learning** is a project aimed at addressing this critical issue.

This project explores the application of **Machine Learning (ML)** techniques to optimize water usage in agricultural irrigation. By leveraging ML algorithms, we aim to predict the precise amount of water required for crops at any given time, considering various factors such as weather conditions, soil type, crop type, and growth stage. This approach not only ensures optimal crop growth but also contributes to sustainable water management.

The objective of this project is to develop a **smart irrigation system** that can autonomously regulate water usage, minimizing waste, and maximizing crop yield. This system could revolutionize farming practices, leading to more sustainable agriculture and a significant reduction in water wastage.

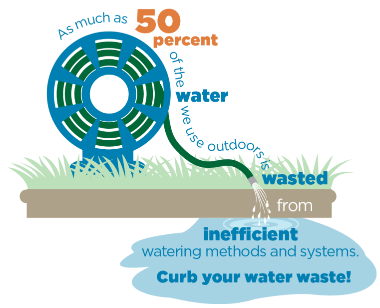


**Fig. 1.** the U.S. Environmental Protection Agency's WaterSense.

In the following sections, we will delve into the details of our proposed system, discussing the ML algorithms used, the data collection process, system design, implementation, and the challenges faced and how we addressed them. We will also present the results of our system in a real-world agricultural setting, demonstrating its effectiveness and potential for broader application. Another project run by the U.S. Environmental Protection Agency's WaterSense program emphasizes the importance of outdoor water efficiency. It provides valuable insights that have been incorporated into the smart irrigation systems.[**?**]

**sustainable water management** in agriculture, paving the way for a future where farming is productive and environmentally responsible.

## 2 Methodology

Our system leverages real-time and historical data to make informed decisions about water management. By analyzing data such as weather forecasts, soil moisture measurements, and historical irrigation patterns, our system can optimize water usage and maximize crop yields. This approach not only conserves water but also contributes to sustainable agriculture practices.

The system uses machine learning algorithms to forecast future water requirements based on real-time data, allowing it to predict the optimal irrigation schedule for Arabidopsis thaliana. This could involve adjusting irrigation schedules based on anticipated weather conditions and continuously monitoring soil moisture levels.

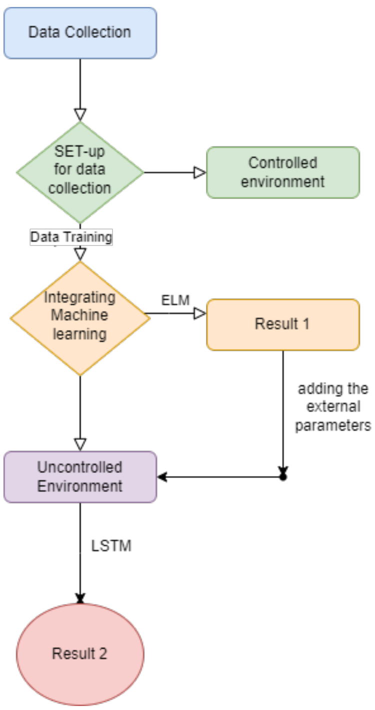


**Fig. 2.** The methodology of the system

### 2.1 Setup and Data collection

Our IoT data collection setup is shown in (Figure 3).Our project setup involves two distinct environments. The first is a controlled environment, where we closely monitor the growth of Arabidopsis thaliana to understand its properties and responses to various irrigation patterns. The second is an uncontrolled environment, which mirrors real-world conditions. Here, we apply our predictive models to manage irrigation effectively, thereby testing the robustness and accuracy of our predictions in the face of unpredictable variables. This dual setup allows us to refine our models continually and ensure their applicability in real-world agricultural scenarios.
Materials used for data collection are;

1. Capacitive soil moisture sensor
2. Connecting wires with diode
3. Logic Level Converter
4. 74LS159 4BIT Decoder
5. ESP8266 based node-MCU

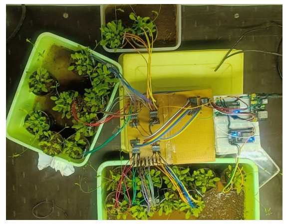

**Fig. 3.** Prototype experiment setup in Green house Long Day Plant Growth facility.

## 3 Soil Profiling & Overview

We Used Combination of **Sandy Clay Loam** & **Spongy Fertilizer** which has negligible effect with actual soil property.
The experiment was framed around the hypothesis that "Developing a predictive irrigation system without reliance on local weather stations is feasible." We aimed at a partial validation of this proposition. Initially, we observed plant behavior within a controlled setting to assess the rate of soil moisture depletion, predominantly characterized by a downward trend. Leveraging this empirical foundation, we proceeded to integrate this information with weather data, as well as pre-existing plant-related metrics.

### 3.1 Machine Learning

The incoming data from sensor input is a time series of data on soil moisture value. This value is standardized for a particular board and sensor-type combination due to varying analog characteristics among components. In the case of uncontrolled environmental irrigation, which our method particularly targets, other time series inputs will be temperature, wind speed, and air humidity data provided by the nearest weather station or online weather forecasting source

We used ELM (Extreme learning machine) to predict future time series points in soil moisture for a controlled environment. ELM is a feedforward neural network a single hidden layer with weights and biases that are randomly generated. In contrast to other neural networks, ELM can attain excellent accuracy without the need for iterative training or bias and weight tweaking. Instead, it computes the output weights depending on the input data using an analytical approach. During the testing phase, the input data is fed into the neural network and the output is generated using the output weights that were previously established.

This method will be applicable in short-duration online prediction of soil moisture in the field while changes in effects of weather parameters are relatively small. Also the increase in moisture after irrigation is instantaneous(Figure 4), so this trend can be disregarded. The following plot in Figure 5 shows an example trend.

For multi-parameter prediction, LSTM has been considered. Other CNN-based approaches might be beneficial, see future work for details. The machine

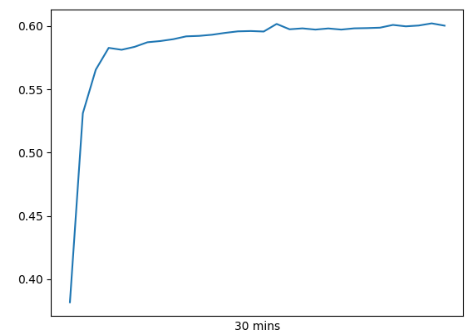


**Fig. 4.** Soil moisture 30 mins after irrigation. $X$ axis is a time interval of 30 minutes, $Y$ axis is relative soil moisture value.

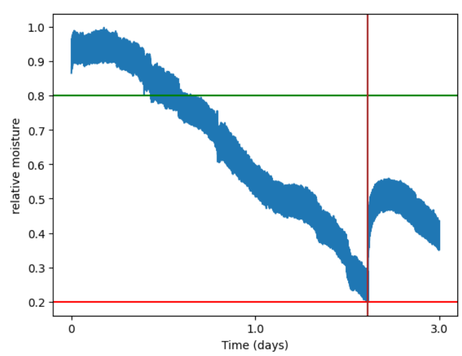


**Fig. 5.** Soil moisture trend after irrigation. The vertical brown line indicates the point of irrigation.

learning algorithms run on standard cloud infrastructure and are lightweight. It can also be run locally with container deployment.The predicted moisture is used for heuristics designed by this project to calculate when to irrigate next. This method primarily tries to keep the predicted moisture in a 'safe' range for the plant. This range can be hard-coded by agronomist advice during sensor calibration or determined automatically from a few days of good initial irrigation schedule by the crop grower. Currently, there is no way to manually start this training anytime while the server is still running but this feature can be added later.

Also this heuristic accounts for events like rainfall that cannot be practically predicted or taken into account for our machine learning methods itself. In case of missed predicted (by weather forecast) rain or other sudden changes that might lead to depletion of moisture below the wilting point of the plant, the user can be alerted to take manual actions.

## 4 Partial Field Trial:

To validate our hypothesis, we did a partial field trial, selecting Arabidopsis as our experimental subject. Subsequent to the initial planting phase, our objective was to observe of the plant's developmental stage from germination to fruition. However, the progress of our study was impeded by the exceedingly elevated outdoor temperatures, which effected the normal growth progression of the specimens. Despite this limitation, we managed to gather a substantial volume of data, sufficient to propel our investigation forward.
We also managed to observe the change of soil moisture without the plant, It also shows us the specific soil property for our experiment.
We run some model & evaluate them **Model Evaluation, KNN / LSTM**

## 5 Results

In ideal scenario,

$$\text{TAW} = \text{field capacity - wilting point of plant}$$
$$\text{RAW} = \text{TAW} \times \text{depletion threshold}$$

Here RAW means Readily available water and TAW is Total available water in the field. Recommended interval for irrigation is $\frac{RAW}{ET_c}$ where $ET_c$ is the evapotranspiration coefficient of the crop calculated as $ET_0\times$ crop coefficient. $ET_0$ is the standard evapotranspiration of the field.

**The performance:**

Our project used both Extreme Learning Machine (ELM) and Long Short-Term Memory (LSTM) models for our data analysis. With its single-layer feedforward neural networks, the ELM model provided us with a quick and efficient way to perform binary classification tasks. However, it lacked the temporal dynamic behavior, which is crucial for time-series data.

On the other hand, the LSTM model, with its unique ability to forget or remember information for long periods, was highly effective for our time-series data. It outperformed the ELM model in terms of accuracy but was computationally more expensive.

To enhance the performance further, we integrated the **K-Nearest Neighbors (KNN)** algorithm with the LSTM model. Combining KNN's simplicity and effectiveness in handling multi-class problems with LSTM's proficiency in processing time-series data significantly improved accuracy. This hybrid approach of KNN+LSTM proved to be a powerful tool for our project, striking a balance between computational efficiency and predictive power.

| **Dataset** | **Accuracy** | **Precision** | **Recall** |
|---|---|---|---|
| Dataset 1 | 0.85 | 0.88 | 0.82 |
| Dataset 2 | 0.89 | 0.91 | 0.87 |
| Dataset 3 | 0.92 | 0.93 | 0.90 |

**Table 1.** Performance of LSTM+KNN model

| | | across time | | |
|---|---|---|---|---|
| | | min | mean | max |
| samples | min | 0.02 | 14.94 | 3.00 |
| | mean | 5.86 | 12.03 | 24.06 |
| | max | 4.8 | 48.00 | 53.70 |

**Fig. 6.** Error in prediction

| **Metric** | **Value** |
|---|---|
| Accuracy | 0.75 |
| Precision | 0.69 |
| Recall | 0.74 |
| F1-Score | 0.75 |

**Table 2.** Performance Matrix for ELM

| Dataset | Accuracy | Precision | Recall |
|---|---|---|---|
| Dataset 1 | 0.85 | 0.88 | 0.82 |
| Dataset 2 | 0.89 | 0.91 | 0.87 |
| Dataset 3 | 0.92 | 0.93 | 0.90 |

**Table 3.** Performance of LSTM+KNN model

**Moisture in uncontrolled environment:**

The scatter plot shows a positive correlation between soil moisture and humidity levels. As the humidity percentage increases, the soil moisture levels also tend to increase. This relationship is crucial for understanding the water dynamics in the soil. It can be used to predict soil moisture levels based on humidity data, which is valuable information for efficient water management. Based on the two images, we can make the following observations and relate them to ourr project on smart water management using a combination of LSTM (Long Short-Term Memory) and KNN (K-Nearest Neighbors) algorithms:

information for efficient water management.

Temperature (°C) and Soil line plot depicts the relationship between temperature and an unknown soil parameter. The plot shows a fluctuating pattern, indicating that temperature changes influence the soil parameter. This relationship could be useful for modeling and predicting soil conditions based on temperature data.

In our smart water management project, the LSTM algorithm can be employed to capture and model the temporal dependencies present in the dataset, which includes features like soil moisture, humidity, temperature, and other relevant variables. The LSTM's ability to learn long-term dependencies makes it suitable for time-series forecasting and sequence prediction tasks.

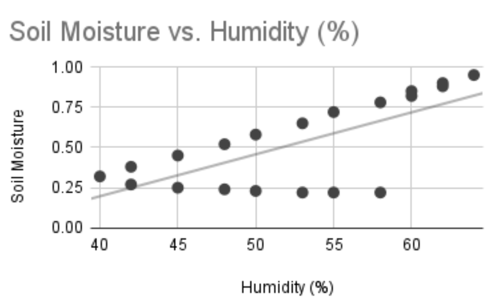


**Fig. 7.** Soil moisture relation with humidity

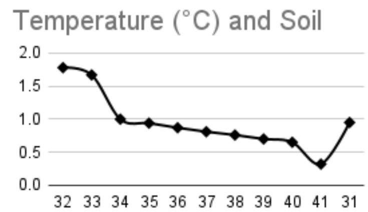


**Fig. 8.** soil moisture with temperature

By combining these two algorithms, we leverage the strengths of both approaches. The LSTM can handle the time-series aspect of the data and capture the temporal patterns, while the KNN can provide an efficient way to classify or predict soil conditions based on the learned patterns from the training data.

## 6 Conclusion

In conclusion, our research underscores the critical role of smart water management, particularly through the integration of machine learning techniques, in revolutionizing traditional irrigation practices. By meticulously collecting and analyzing data in controlled environments, we have unveiled the potential of predictive irrigation in optimizing water usage and improving agricultural efficiency.

Transitioning from controlled experiments to uncontrolled environments, we meticulously assessed the performance of our predictive models using a range of tools. The observed predictive capabilities further validate the efficacy of our approach in real-world settings.

Our findings emphasize the urgency of embracing innovative methodologies to tackle the pressing challenges associated with conventional smart irrigation techniques (where weather station is essential). Through the continued refinement of predictive models and their deployment in practical scenarios, we anticipate a transformative shift towards more sustainable water resource management in agriculture.

Crucially, our research demonstrates the feasibility of developing crop-specific predictive models without reliance on local weather stations. This highlights the scalability and accessibility of predictive irrigation systems, paving the way for widespread adoption and implementation.

As we forge ahead, our commitment remains steadfast in advancing the frontier of predictive irrigation technology. By harnessing the power of machine learning, we aim to catalyze a resilient and productive agricultural landscape, ensuring a brighter future .