SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage: https://jaraset.com/
ISSN: 2462-1943

DEVELOPMENT OF A MACHINE LEARNING-BASED SMART IRRIGATION SYSTEM FOR OPTIMAL WATER RESOURCE MANAGEMENT

Amit Solanki [1*], Dr. Atul Gandhi [2], Prof. (Dr.) Vinesh Agrawal [3]

[1] *Research Scholar, Department of Electrical Engineering, Sangam University, Rajasthan, India*
[2] *Assistant Professor, Department of Electrical Engineering, Sangam University, Rajasthan, India*
[3] *Professor, Department of Electrical Engineering, Sangam University, Rajasthan, India*

ARTICLE INFO

Article history:
Received: 08-10-2024
Received in revised form: 27-10-2024
Accepted: 09-11-2024
Available online: 08-05-2025

Keywords:

Smart Irrigation, Irrigation Automation, Machine Learning, Water Conservation, Food Security, Water Management Plan.

ABSTRACT

Global warming and climate change are significant contributors to water scarcity, necessitating the efficient management of water resources for long-term sustainability. Agriculture, as one of the largest consumers of water, plays a critical role in this scenario. However, the water used in agricultural activities is often polluted and unsuitable for reuse. Thus, optimizing water management systems in agricultural irrigation is essential. This paper introduces an automated crop irrigation system that leverages real-time soil sensor readings to predict water treatment plans. Additionally, the system integrates weather conditions into its decision-making process before initiating water supply. A decision-making function has been developed to predict water treatment requirements and future weather conditions accurately. To achieve this, an Artificial Neural Network (ANN) algorithm has been trained to perform dual predictions.

To enhance the effectiveness of the dataset used for training the machine learning model, a preprocessing algorithm for soil moisture sensor data has been proposed. Experiments and simulations were conducted to evaluate the system's performance. The results demonstrate that the proposed ANN-based weather prediction technique achieves a higher accuracy of 99.4% compared to Support Vector Machine (SVM), which achieves 95% accuracy. For soil condition predictions, the system delivers an accuracy of 88.4%. Moreover, the system's decision-making and training times were assessed, showing that decisions can be made within fractions of a second. The training process, conducted on location-specific data, requires minimal time and only needs to be performed once for a given location. These findings highlight the efficiency and practicality of the proposed automated irrigation system in optimizing water management in agriculture.

Introduction

The machine learning (ML) techniques can be used in various real world problems. Mainly these techniques have employed in order to perform prediction, classification and decision making. The prediction aims to approximate a continuous value, classification perform categorization, and decision making capture and identify key insights to react on a specific event [1]. In this presented work, the ML techniques are being used for decision making task. In precision farming it is essential to measure the requirement of crops and supply appropriate amount of water to crops for improved production and quality [2].

In traditional farming process, crop irrigation decisions are highly depends on the soil conditions such as temperature and moisture. In addition, the weather conditions are also influencing variable for accurate irrigation decision making process [3].

The aim is to involve the analysis and utilization of weather information (prediction) in order to make more precise and practical decision making for crop practical decision making for crop irrigation. Therefore:

1. First explore the sensor reading dataset to correctly identify the type of prediction or classification problem.
2. Next, explore the techniques of weather data collection, analysis and utilization.
3. Third, introduce an algorithm to combine weather prediction information into the irrigation system automation.

The proposed model is trying to automate decision making and water treatment practice for more practical use. The proposed technique maps the irrigation decisions based on the relationship between the soil temperature, moisture and weather conditions. Therefore, the data exploration is the essential task before proposing any method. After exploration we can decide the appropriate ML algorithm for employing with the proposed system. This section presents an overview of the work involved in this paper, next section provide a detailed investigation of the utilized dataset.

Data Collection and Exploration

In order to explore the data nature and to identify the type of problem hidden in the sensor readings analysis we utilize a predefined dataset. Thus, a Soil Moisture and Temperature - Data set is downloaded from the KBS LTER core database [4]. The dataset is consisting of sensor readings and labelled with appropriate water treatment labels.

Table 1: KBSLTER core database attributes 1

Variate	**Description**	**Units**
Year	Yearin which sampling date occurs	
Sample date	Sampling date	
Trt	Treatments	
T0-2cm	Temperature,0 to 2cm	Celsius
T3-15cm	Temperature, 3 to 15cm	Celsius

W0-2CM	Water content 0 to 2cm	Cubic centimeters \cubic centimeters
W3-13CM	Water content 3 to 15cm	Cubic centimeters \cubic centimeters

The dataset is the collection of soil moisture and temperature readings for 10 years, each day. It contains a total of 15706 readings. The attribute "Treatment_Type" has needed to predict. There are five treatment types (i.e. 1, 2, 4, 6, 7 and 8). Figure 2 shows the raw dataset samples.

Table 2: Raw dataset samples

	Type	**T0-2cm**	**T3-15cm**	**W0-2cm**	**W3-15cm**
0	1	1.68	1.77	0.02	0.18
1	2	1.77	1.87	0.02	0.19
2	7&8	1.63	1.72	0.02	0.18
3	1	1.65	1.73	0.02	0.18
4	2	1.76	1.85	0.02	0.19

The treatment type or classes are first encoded to(0, 1, 2, 3, and 4) using a label encoder function. During this we want to check is the dataset has any class imbalance problem or not [5]. Therefore, a count plot has been prepared to know the distribution of samples into the water treatment type. Figure 3 shows the class distribution of the water treatment dataset.

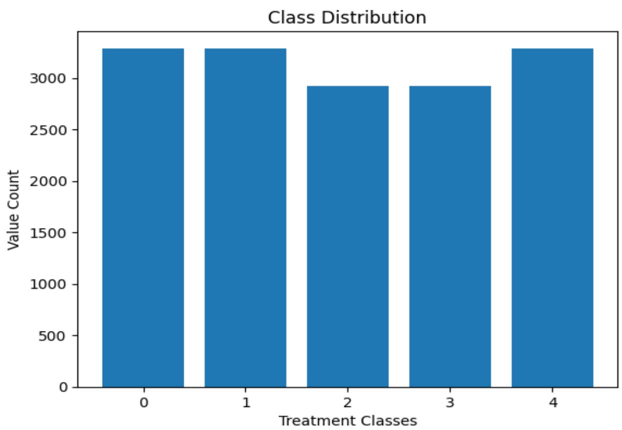

Fig 1: water treatment class distribution

Next, we focus on the dataset, which consist of seven attributes; among them two columns is the time stamp. Therefore, initially we consider the data and problem as time series problem

[6]. In this context, we eliminate the “Years” attribute and “Date” attribute is converted into the index column.

Table 3: Final dataset used for analysis

Treatment type	**T0-2cm**	**T3-15cm**	**W0-2cm**	**W3-15cm**
0	1.68	1.45	1.36	1.58
1	1.77	1.85	1.25	1.25
4	1.63	1.63	1.63	1.45
0	1.65	1.56	1.69	0.28
1	1.76	1.55	1.59	0.54
4	1.60	1.35	1.42	0.36
0	1.63	1.57	1.33	0.57
1	1.75	1.69	1.54	0.24
4	1.59	1.59	1.27	1.28
0	1.59	1.35	1.54	1.33s

Additionally, the treatment type of these readings is different from each other therefore, if we apply any time series operation or time series based pre-processing then data loss has been possible. Therefore, we drop the idea of analyzing the water treatment prediction problem as time series forecasting problem [7].

Thus, we consider this problem as the classification or pattern recognition problem [8]. In this context, we drop time information from the dataset to avoid over-fitting and under-fitting problem [9]. Next, we are investigating the influence of attributes over the classification. Therefore, the entire dataset has grouped according to the treatment type and prepare a line plot to understand the patterns of the data attributes for making different set of decisions, the temperature attribute’s line plot is given, which shows high over-lapping between them. Additionally, the attributes are following a cyclic trend in temperature.

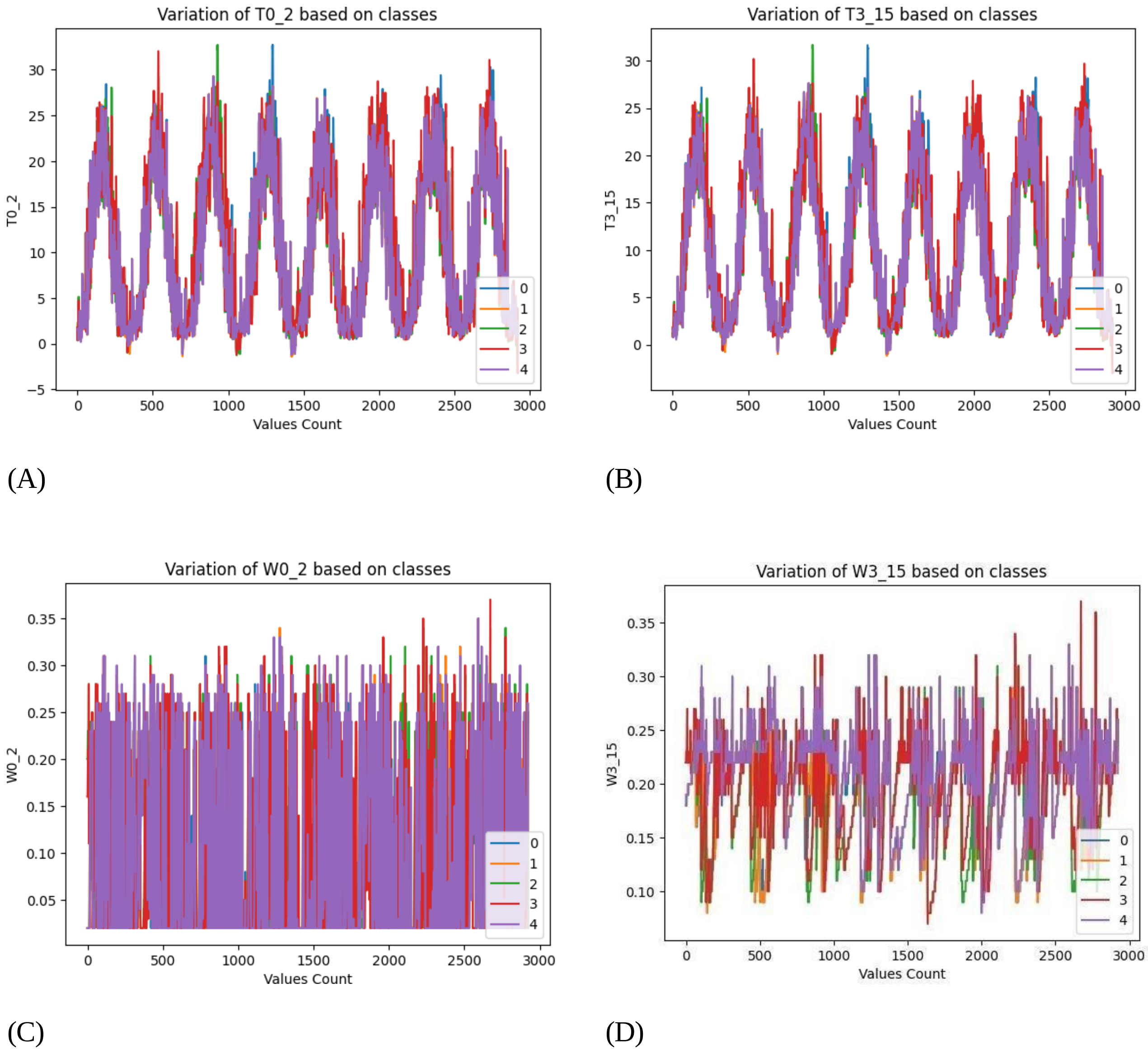

Fig 2: Attributes plot according to the treatment type

The dataset contains a time stamp attribute of the time series problems are configured to perform regression but the discussed problem is a multi-class classification problem [10]. The second issue is belongs to the frequency of treatment, and the type of treatment [11].

Table 4: Example of data samples

Treatment type	**T0-2cm**	**T3-15cm**	**W0-2cm**	**W3-15cm**
0	1.68	0.25	1.35	1.52
1	1.66	1.26	0.25	0.26
4	1.25	1.25	0.14	0.43
0	0.25	0.25	1.25	0.56

1	0.75	0.48	2.25	0.25

In this context, are proposed to pre-process the data for making it effective for learning algorithm. The pre- processing algorithm follows the following steps:

Table 5: Pre-processing Algorithm

Input: Dataset D

Output: pre-processed Dataset P

Process:

1. $S = D.SortValues(byDate)$
2. $G = S.GroupBy(Date)$
3. $foreach g_i in G$
   a. //where $g_i = \{d_1, d_2, \ldots d_n\} instances$
   b. $M = \frac{1}{n}\sum_{j=1}^{n} d_j$
   c. $P.Add(M)$
4. End for
5. Return P

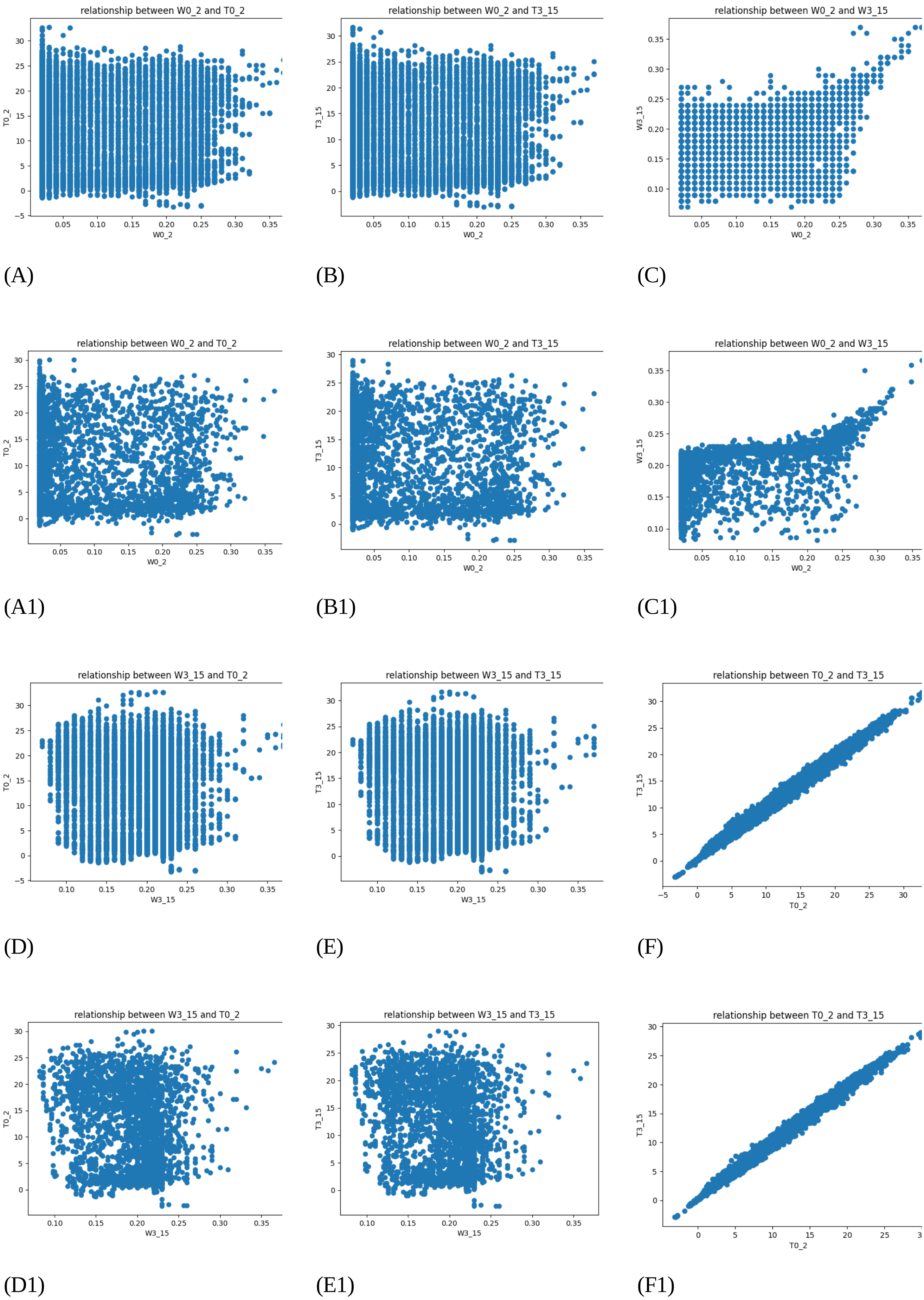

Fig 3: Relationship between dataset attributes before and after applying the pre-processing algorithm

Weather Forecasting

The weather data and future trend prediction is now in these available by using different Application Programming Interface (API). There are two popular APIs available for gathering location aware weather information. According to the recorded performance of the ANN algorithms we found the accuracy of the algorithm has remained constant and results to 88.4% accuracy for both the algorithms. In addition, the SVM classifier has soused and the comparison have been done for both the conditions before and after pre-processing of the data.

Open Weather Map

It is a service for weather data, which includes current weather, forecasts, and historical data for developers. This API enables the utilization for web and mobile applications. API also supports JSON, XML, and HTML endpoints. A free and limited version of API is available. Additionally for more than 60 requests per minute paid subscription required. To use this API, we need an API key [13].

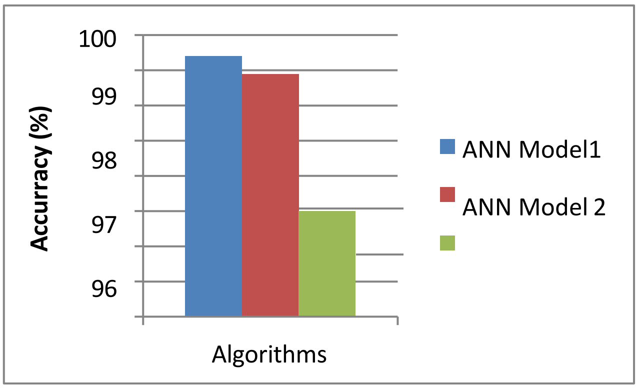

Fig 4: Accuracy for weather prediction

In these experimental scenarios in place of collecting data directly from farmland by using sensors we have utilized a temperature and moisture sensor reading dataset. In addition, in place of taking weather data from API or any other source, we have considered the weather dataset obtained from Kaggle. The detailed study about the dataset has been

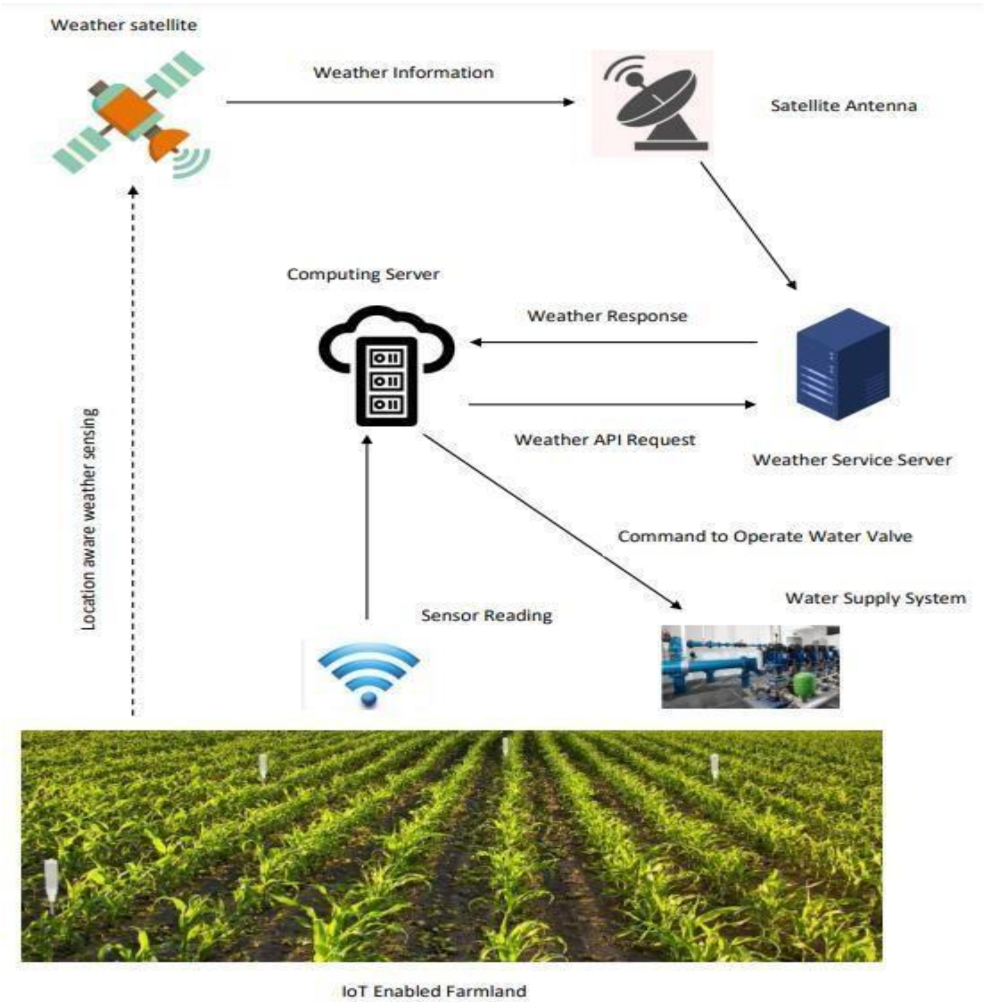

Fig 5: shows the system architecture to configure both the components weather and water treatment prediction

Data Collection

In this work, the problem of water irrigation has considered thus the data can be collected from the farm land. In recent discussed in previous section. In next section, we are discussing how the computational server has utilizing the weather and sensor reading to predict the water treatment plan automatically.

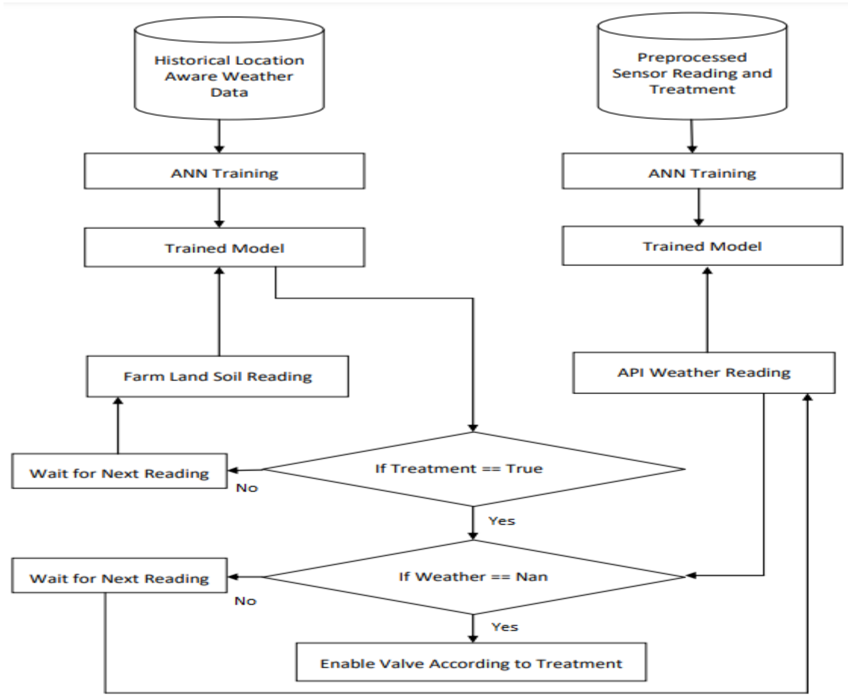

Fig 6: Process of utilizing the weather forecast and soil moisture readings in automation of irrigation system

Data Processing

The steps of utilizing the weather and sensor reading in the proposed proactive management system have been demonstrated in this section. The computational system work flow for obtaining the accurate and automated water supply decision making processing. During the investigation we have implemented and experimented with the two versions of ANN algorithm's configuration and SVM algorithm for selecting appropriate prediction technique. The considered prediction algorithm provides 88.4% accurate prediction for soil moisture and temperature sensors additionally producing 99.4% accurate results for weather prediction task. However, the system required to train only once for initialization of prediction algorithm. The training time of the selected ANN algorithm and SVM algorithm is also measured. The training time of the algorithm has been given in terms of seconds (sec). According to the training time of the algorithms the weather data takes higher time due to large number of parameters to learn as compared to soil moisture dataset.

Conclusions

The agricultural use of fresh water is never recovered again. Additionally, recycling needs a significant amount of time and artificial process are expensive. Therefore, fresh conservation is a challenging task in agriculture. In this paper, we are aimed to minimize the utilization of water without affecting the crop yield. Therefore, an automated water treatment system has been proposed. Additionally, it various components has been discussed. The model consists of three main parts data collection, data analysis and decision making. First, we discussed how the data can be collected from farm land by using the sensor hard ware which is influenced by our recent work [16].After collecting the localized sensor readings the pre-trained ANN algorithm has been used to decide water treatment decisions. These sensor readings are indicating the water requirements of the crop or farmland. Next the weather conditions are also influencing the water treatment decisions. Therefore, a weather forecasting system is also incorporated to predict the possible weather conditions.

Finally, a decision function has been introduced to make decisions of water supply enabling and disabling. This function accepts the predictive outcomes of the ANN algorithms. For weather conditions ANN perform 99.4% accurate predictions and for soil condition before data pre-processing is predicts only 21% correct treatment.

References

1. I. H. Sarker, “Machine Learning: Algorithms, Real- World Applications and Research Directions”, SN Computer Science volume 2, Article number: 160 (2021
2. A. Monteiro, S. Santos, P. Gonçalves, “Precision Agriculture for Crop and Livestock Farming Brief Review”, Animals 2021, 11, 234.
3. M. Dhanaraju, P. Chenniappan, K. Ramalingam, S. Pazhanivelan, R. Kaliaperumal, “Smart Farming: Internet of Things (IoT)-Based Sustainable Agriculture”, Agriculture 2022, 12(10), 1745
4. S. Sarkar, A. Pramanik, J. Maiti, G. Reniers, “Predicting and analyzing injury severity: A machine learning-based approach using class-imbalanced proactive and reactive data”,

Safety Science 125 (2020) 104616

5. A. Zeroual, F. Harrou, A. Dairi, Y. Sun, "Deep learning methods for forecasting COVID-19time-Series data: A Comparative study", Chaos, Solitons and Fractals 140 (2020) 110121

6. M. Saeed, "A Guide to Obtaining Time Series Datasets in Python", in Python for Machine Learning, March29, 2022, https://machinelearningmastery.com/a-guide-to- obtaining-time-series-datasets-in-python/

7. EzioPreatoni, Stefano Nodari, Nicola Francesco Lopomo, "Supervised Machine Learning Applied to Wearable Sensor Data Can Accurately Classify Functional Fitness Exercises Within a Continuous Workout", Front. Bioeng. Biotechnol. 8:664, July2020

8. X. Ying, "An Overview of Overfitting and its Solutions", IOPConf. Series: Journal of Physics: Conf. Series1168(2019)022022

9. S. Ruberto, V. Terragni, J. H. Moore, "Towards Effective GP Multi-Class Classification Based on Dynamic Targets", GECCO '21, July 10–14, 2021, Lille, France, ACM

10. J. Bernard, T. Ruppert, O. Goroll, T. May, and J. Kohlhammer, "Visual-Interactive Preprocessing of Time Series Data", SIGRAD 2012

11. D. C.T. Pérez, J. R. Reséndiz, R. A. G. Loenzo, J. C. J. Correa, "Support Vector Machine-Based EMG Signal ClassificationTechniques:AReview",Appl.Sci.2019, 9, 4402 open weathermap.or

12. M. Breuss,"Beautiful Soup: Builda WebScraper With Python", https://realpython.com/beautiful-soup-web- scraper-python/ https://www.kaggle.com/datasets/ananthr1/weather- prediction

13. P. Pandey, S. Agarwal, "A Low Cost Smart Irrigation Planning Based on Machine Learning and Internet of Things", CurrAgri Res 2023; 11(2).

14. Rajendra, K. ., Subramanian, S. ., Karthik, N. ., Naveenkumar, K. ., &Ganesan, S.. (2023). Grey Wolf Optimizer and Cuckoo Search Algorithm for Electric Power System State Estimation with Load Uncertainty and False Data. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2s), 59–67.

https://doi.org/10.17762/ijritcc.v11i2s.6029

15. Kshirsagar, P.R., Reddy, D.H., Dhingra, M., Dhabliya, D., Gupta, A. Detection of Liver Disease Using Machine Learning Approach (2022) Proceedings of5th International Conference on Contemporary Computing and Informatics, IC3I 2022, pp. 1824-1829.

16. Martínez, L., Milić, M., Popova, E., Smit, S., & Goldberg, R. Machine Learning Approaches for Human Activity Recognition. Kuwait Journal of Machine Learning, 1(4). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/146