BINDURA UNIVERSITY OF SCIENCE EDUCATION

FACULTY OF SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE

IOT-BASED SMART IRRIGATION SYSTEM USING MACHINE LEARNING

BY

KAPIMBI MELISA

B200772B

SUPERVISOR: MR H CHIKWIRIRO

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE BACHELOR OF SCIENCE HONOURS DEGREE IN

COMPUTER SCIENCE

ABSTRACT

The efficient management of water resources in agriculture is critical for sustainable farming practices. This research project presents an innovative IoT-based smart irrigation system that leverages machine learning to optimize irrigation water needs. The system utilizes real-time data from humidity, temperature, and soil moisture sensors deployed in the field to predict water requirements. An Arduino Uno microcontroller acts as the brain of the system, collecting data from the sensors and transmitting it to the internet via a WIFI module. The collected data is then sent to a random forest algorithm, which predicts water needs and triggers the pump to automatically open or close. The system features a mobile application and web interface, providing farmers with real-time sensor data and pump status. Remote monitoring and automation enable efficient water usage, reduced energy consumption, and improved crop health. The results show a significant reduction in water usage and improved crop yields, making this system a valuable contribution to the development of sustainable agriculture practices. This research contributes to enhancing the accuracy and reliability of irrigation water demand forecasting. The developed model offers valuable insights for farmers, aiding in better water resource management and sustainable agricultural practices. The findings highlight the potential of machine learning in optimizing irrigation practices, leading to improved resource allocation and reduced water wastage in the agricultural sector.

APPROVAL FORM

The undersigned certify that they supervised (student) Melisa T. Kapimbi's (Registration Number: B200772B) dissertation entitled IOT BASED SMART IRRIGATION SYSTEM USING MACHINE LEARNING, submitted in partial fulfillment of the requirements for the Bachelor of Computer Science Honors Degree of the Bindura University of Science Education.

……………………..	……………………….	…………………….
STUDENT	**SIGNATURE**	**DATE**
……………………..	………………………..	…………………….
SUPERVISOR	**SIGNATURE**	**DATE**
……………………..	…………………………..	……………………..
CHAIRPERSON	**SIGNATURE**	**DATE**
……………………………	…………………………..	……………………..
EXTERNAL EXAMINER	**SIGNATURE**	**DATE**

ACKNOWLEDGEMENTS

The success of this project is a collective effort, and I am unable to attribute its accomplishments to any one individual. I would like to express my deepest gratitude to the Almighty God for His infinite grace and blessings. I also extend my sincere appreciation to all authors, both acknowledged and unacknowledged, whose contributions have shaped the final outcome. I am particularly indebted to my supervisor, Mr. Chikwiriro, for his expert guidance, unwavering kindness, and motivational support throughout my research journey. Additionally, I would like to thank my classmates, friends, and family for their unrelenting support and encouragement. Their collective efforts have been invaluable to me.

DEDICATION

I dedicate this to God and my family.

TABLE OF CONTENTS

Table of Contents

LIST OF ACRONYMS

loT – Internet of Things

LCD – Liquid Crystal Display

AI - Artificial Intelligence

ML – Machine Learning

GPS – Global Positioning System

IDE – Integrated Development Environment

FAO- Food and Agriculture Organization

ARDA - Agricultural Rural Development Authority

MLR -Multiple Linear Regression

SVM - Support Vector Machine

WFP - World Food Programme

GDP - Gross Domestic Product

ZINWA - Zimbabwe National Water Authority

TP – True Positive

TN – True Negative

FP – False Positive

FN – False Negative

LIST OF FIGURES

CHAPTER 1: PROBLEM IDENTIFICATION

1.1 Introduction

The way we interact with and maintain physical systems has changed as a result of the internet of things (IoT). It provides remote monitoring and control of physical settings by combining wireless sensor networks with internet-based technologies. IoT enables seamless observation and analysis of physical system activity, using sensor data to model, forecast, and automate numerous operations, altering our understanding and optimization of physical infrastructure. According to Mokyr et al. (2015), the early twenty-first century saw a fast acceleration in technological innovation, resulting in new paradigms and techniques that have profoundly affected the development of the Internet as we know it today. The goal of the new idea known as the Internet of Things (IoT) is to use the Internet to link a variety of consumer electronics. More convenient and effective ways of living have been made possible by the exponential growth in the number of internet-connected gadgets over time. The advent of smart household appliances, such as microwaves, refrigerators, and lighting systems, which are simple to connect to and control from a central hub like a smartphone, is a good example of this trend. This connection enables users to easily modify the operation of these associated devices remotely, a huge breakthrough in everyday life made possible by the growing IoT ecosystem. By 2025, it is expected that there will be up to 7 trillion "connected devices" or "Things" linked to the internet. The spread of IoT devices is predicted to greatly improve people's daily lives. The Internet of Things movement has had a far-reaching impact, having been commercialized and integrated into a variety of areas such as business, the military, medical, and agriculture. The widespread acceptance demonstrates how the interconnection afforded by the internet of things is affecting not only individual lifestyle, but also driving innovation and optimization across a wide range of industries.

The emergence of the Internet of Things (IoT) is rapidly transforming a wide range of fields, including precision agriculture, control systems, data science, and artificial intelligence. In the domain of precision agriculture, IoT technologies are leveraged to provide targeted, optimized inputs for healthy plant growth and improved productivity. IoT-enabled agriculture utilizes sensor networks to measure and monitor a diverse array of agricultural variables, such as plant growth, rainfall, soil moisture, temperature, pH, humidity, airflow, soil properties, and solar radiation. The data collected from these sensors is then mined and analyzed to enable applications like soil erosion control, plant and animal disease diagnosis, precision fertilizer application, comprehensive field monitoring, crop yield forecasting, and smart irrigation systems. This data-driven approach facilitated by IoT is revolutionizing the way agriculture is practiced, leading to greater efficiency, sustainability, and productivity. Irrigation has long been used to provide steady agricultural production despite seasonal changes, allowing food supply to keep up with the world's rising population. The anticipated expansion in global population has increased the emphasis on methods and practices that can boost agricultural output while ensuring food security. In this context, research into IoT-powered irrigation systems is getting significant traction. Smart irrigation solutions provided by the Internet of Things could increase crop yields, enhance water conservation, minimize labour requirements, and provide farmers with a more efficient and appealing irrigation method. By using IoT capabilities, these smart irrigation systems have the potential to address the pressing challenge of meeting future food demands through sustainable, data-driven farming methods.

Furthermore, one of the most transformational technologies influencing the future of farming is machine learning. By leveraging the power of algorithms and data analysis, machine learning is revolutionizing various facets of agriculture, from crop management to resource optimization. Machine learning plays a pivotal role in precision agriculture, a paradigm shift from traditional farming practices. Through the integration of sensors, drones, and satellite imagery, machine learning algorithms analyze vast datasets to provide farmers with insights into crop health, soil conditions, and optimal harvesting times. This data-driven approach allows for targeted interventions, minimizing resource usage while maximizing yields. The Random Forest algorithm, with its capacity to handle complex and diverse datasets, emerges as a transformative force in modern agriculture. Its applications, ranging from predicting crop yields to managing pests and optimizing resource usage, signify a shift towards data-driven decision-making in farming practices. As we navigate the complexities of an ever-changing agricultural landscape, the Random Forest algorithm stands as a beacon of innovation, offering sustainable solutions, and paving the way for a more efficient and resilient future in agriculture.

1.2 Background to the study

Every developed nation's foundation is agriculture. It consumes 85% of available freshwater resources worldwide, and this proportion remains dominating in water use due to population expansion and growing food demand. Consequently, effective water management is given top priority in many cropping systems in semi-arid and arid areas. To optimize agricultural water utilization, an automated irrigation system is required. The purpose of an automated irrigation system is to prevent over- and under-irrigation. Over-irrigation is caused by improper wastewater distribution or management, as well as chemicals that pollute the water. Under irrigation, soil salinity increases, resulting in the accumulation of harmful salts on the soil surface in places with significant evaporation.

In a world where water is becoming increasingly scarce, its usage must be carefully optimized. In arid and semiarid environments, irrigation is critical to agricultural output. However, in many countries and regions, limited water resources have become 95% of the bottleneck of regional economic and social development (Sharma, 2021). IoT-driven smart irrigation systems offer a promising solution by optimizing water usage, reducing labor, and enhancing crop yields. Irrigated agricultural systems, which consume the majority of the world's freshwater, must improve their management and conserve as much of this limited resource as possible.

Zimbabwe is an agricultural country where irrigation plays a significant role in food production (Mubvami et al., 2020 Recent years have seen a deterioration in Zimbabwe's agricultural industry due to mismanagement of resources, drought, and a lack of knowledge among farmers. Due to this situation, the once-thriving agricultural sector has taken a dark turn, leading to food shortages and the government's need to rely on outside assistance and purchase crops from nations like Brazil, which leads the world in wheat production. A more concentrated effort must be made to connect the local farming industry with new and developing technology that would increase farming efficiency if Zimbabwe is to regain its position in the global agricultural sector.

1.3 Statement of the Problem

Upon thorough reflection on the agricultural sector in our country, the researcher has identified multifaceted challenges that demand innovative solutions, with a particular focus on harnessing the capabilities of the Internet of Things (IoT) and machine learning, specifically the random forest algorithm. Agriculture is important to Zimbabwe's economy, with the bulk of farmers producing crops using rain-fed methods. However, with climate change leading to increased variability in rainfall patterns, there is a growing need for more efficient irrigation practices (Ndengu,2019). The lack of accurate and timely information on water demand has led to inefficient use of irrigation resources and low crop yields. This problem is exacerbated by the unpredictable rainfall patterns and frequent droughts that Zimbabwe experiences. The Food and Agriculture Organization (FAO) reports that the country's agricultural output has been dropping in recent years because of these issues. Therefore, there is a pressing need to develop effective strategies to predict irrigation in Zimbabwe, to optimize the use of water resources and improve crop yields.

1.4 Research Objectives

1. To analyze the applicability of the random forest algorithm and IoT in monitoring and controlling the irrigation process.

2. To design and implement an IoT smart irrigation system that integrates soil moisture, humidity, and temperature sensors, and utilizes the random forest algorithm for accurate prediction of irrigation needs

3. To assess the effectiveness of utilizing the random forest algorithm and IoT in agriculture for improved water management.

1.5 Research questions

1 How can the analysis of utilizing the random forest algorithm and IoT for predicting irrigation water demands be conducted effectively?

2 What are the main considerations and procedures in designing and implementing an IoT smart irrigation system that uses the random forest algorithm to accurately estimate irrigation water demands?

3 How can the effectiveness of the random forest algorithm and IoT in improving agricultural practices and water management be evaluated reliably and comprehensively?

4 How does the integration of IoT and the random forest algorithm enhance the accuracy and efficiency of water management in smart irrigation systems?

5 What are the major parameters and data inputs required by the random forest algorithm to anticipate irrigation needs in smart irrigation systems?

6 What are the potential challenges and limitations of implementing smart irrigation systems with IoT and the random forest algorithm, and how can they be addressed?

7 How do smart irrigation systems that use IoT and the random forest algorithm compared to traditional irrigation methods in terms of water conservation and crop yield?

1.6 Research propositions/hypothesis

H_0: Implementing an IoT-based smart irrigation system, integrated with the random forest algorithm, will lead to a significant reduction in water wastage, enhance crop health, and optimize agricultural productivity.

H_1: Implementing an IoT-based smart irrigation system, integrated with the random forest algorithm, will not lead to a significant reduction in water wastage, enhanced crop health, and will not optimize agricultural productivity.

1.7 Justification/significance of the study

The study is of significant importance, especially in a country like Zimbabwe, where the agricultural sector is a critical component of the economy. The application of machine learning algorithm can help farmers and irrigation authorities to accurately predict irrigation water demand, thereby optimizing water usage and increasing crop yield. Zimbabwe has experienced prolonged droughts in recent years, resulting in reduced crop production and food insecurity. Thus, this study's findings can help improve food security in the country.

Moreover, the study's application of machine learning can also contribute to the development of a more sustainable and efficient irrigation system. By accurately predicting irrigation water demand, farmers can avoid over-irrigation, which can lead to water logging, soil degradation, and reduced crop yield. Additionally, the optimization of water usage can lead to reduced water wastage and increased water-use efficiency. These outcomes are critical in a country like Zimbabwe, where water resources are scarce, and the need to manage them sustainably is of utmost importance.

1.8 Assumptions

- The sensors used in the system will provide accurate and consistent data.
- The algorithm will be able to accurately interpret and use the data from the sensors.
- The network connectivity for the system will be reliable.

1.9 Limitations/challenges

- availability and quality of data used to train the algorithm

- The system relies on a network connection to transmit data, and any connectivity issues could impact the system's ability to make predictions.
- Because of economic uncertainty, there might be alterations to the budget allocated for the research.

1.10 Scope/delimitation of the research

The scope of this research encompasses the design, implementation, and evaluation of an IoT-based smart irrigation prototype. The prototype incorporates soil moisture sensors, humidity sensors, and temperature sensors to collect real-time data from the farm. The random forest algorithm is employed to analyze this data and predict optimal irrigation conditions. The web interface, Thing Speak, allows farmers remote access for monitoring and controlling irrigation processes.

1.11 Definition of terms

Irrigation - is the delivery of water to land or crops to aid growth, usually through channels.

Water demand - is the quantity of water required by customers to meet their demands.

Machine learning - is the usage and development of computer systems that can learn and adapt without explicit instructions, relying on algorithms and statistical models to analyze and derive conclusions from data patterns.

Precision agriculture - is a cutting-edge farming method that makes use of data analytics, GPS, and sensors to optimize the use of resources including water, fertilizer, and pesticides.The goal is to increase efficiency, reduce waste, and maximize crop output by making data-driven decisions based on unique field circumstances.

Algorithms - are processes or sets of rules that must be followed while performing calculations or other problem-solving tasks, particularly by computers.

1.12 Conclusion

In conclusion, Chapter 1 prepares the groundwork for investigating the revolutionary potential of an IoT-based smart irrigation system that employs the Random Forest algorithm to address significant agricultural difficulties in Zimbabwe. The chapter provides a comprehensive overview of the significance of incorporating technology, particularly IoT and machine learning, to tackle water scarcity, and climate change, and optimize crop yields. The research problem is clearly defined, emphasizing the need for data-driven irrigation strategies. The research objectives and questions guide the study towards a detailed investigation. While assumptions are made about sensor accuracy and algorithm interpretation, potential limitations such as data quality, network reliability, and budget uncertainties are acknowledged. The study's significance lies in its potential to rejuvenate Zimbabwe's agricultural sector and contribute to global trends in smart agriculture.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

The researcher outlines previous and present systems that have been researched by other writers and are relevant to the current study topic in this chapter, in addition to addressing the research concerns. This will be highly beneficial to the author because it serves as a guide for locating solutions, strategies, and methods that previous authors have employed to address research problems in the past. Okoli (2015) states that a comprehensive and critical study of all significant previous research on a given topic must be included in a literature review. A literature review shows what has been done in the past, where there are knowledge gaps, and what issues still need to be tackled.

2.2 Irrigation

Irrigation is the artificial application of water to soil, through a variety of pipe, dam, and canal systems. The use of irrigation helps with landscape upkeep, agricultural crop growth, and wasteland reclamation. In places where rainfall is infrequent or inconsistent, irrigation can be used to either fully replace or enhance natural rainfall. Many agricultural systems depend heavily on irrigation to make sure that crops have the water they require to develop and flourish (FAO, n.d.). Farmers can boost agricultural yields, enhance crop quality, and lower the chance of crop failure from drought by providing enough irrigation. Furthermore, irrigation expands the amount of land that may be used for cultivation by enabling farmers to raise crops in areas with limited rainfall (Islam et al., 2019). Irrigation is essential to agriculture overall because it helps farmers increase food production, enhance food security, and promote sustainable agricultural development.

2.2.1 Irrigation in Zimbabwe

With 60–70% of the population relying on agriculture for a living, Zimbabwe is primarily an agricultural nation. About 30% of Zimbabwe's total agricultural production is attributed to irrigation, making it a vital component of its agricultural sector. Just 50% of the estimated 270,000 hectares of arable land in the nation are currently under irrigation. Numerous factors contribute to the low utilization rate, such as restricted water access, poor financing for irrigation development, and a lack of infrastructure.

A number of boards and organizations have been established to assist and encourage the construction of irrigation infrastructure in Zimbabwe in an effort to improve irrigation. The Zimbabwe National Water Authority (ZINWA) is one such agency that oversees water resources and supplies farmers with water for irrigation. Another organization involved in promoting irrigation in Zimbabwe is the Agricultural Rural Development Authority (ARDA), which is responsible for managing state-owned farms and promoting agricultural development in the country. ARDA is working on a number of irrigation projects to give smallholder farmers access to irrigation water in an effort to increase their production. Three irrigation schemes are among these projects: the Muzarabani, Wenimbi, and Chisumbanje schemes.

Apart from these groups, the government of Zimbabwe has also formulated policies with the objective of encouraging the growth of irrigation in the nation. In order to encourage sustainable irrigation growth, the National Irrigation Policy was created in 2016. It does this by offering a framework for the construction of irrigation infrastructure, enhancing the management of water resources, and encouraging private sector participation in irrigation development.

The technologies that are currently being used for smallholder irrigation in Zimbabwe, including drip irrigation, treadle pumps, and bucket and rope irrigation, were highlighted by Marinda (2016). The author discovered that better technologies are required, ones that are more suitable for Zimbabwe's agroecological circumstances and smallholder farmers.

In Zimbabwe's smallholder agricultural communities, solar-powered drip irrigation systems are another technological use (Ilemobade, 2017). According to the author, these methods have the ability to lower labor costs and water usage while simultaneously increasing yields and earnings. They did point out that there are certain obstacles to the uptake and durability of these systems, such as expensive initial expenses and a dearth of professional assistance.

2.2.2 Irrigation Water Shortages In Zimbabwe

Irrigation water scarcity has a substantial impact on agricultural production in Zimbabwe, as agriculture is the mainstay of the economy. According to the Zimbabwe National Water Authority, only 165,000 hectares of the 1.5 million hectares of arable land in the country are currently under irrigation, with the majority relying on rainfall. The lack of adequate irrigation infrastructure and water management has led to frequent water shortages, resulting in reduced crop yields and a negative impact on food security (Lawrence et al., 2016). For example, during the 2019-2020 agricultural season, the grain harvest in Zimbabwe was estimated at 907,628 tonnes, which was 33% lower than the previous season due to drought and poor irrigation.

Shortages of irrigation water have an effect on more than just agriculture and food security. The agricultural industry employs about 67% of the workforce and accounts for about 17% of Zimbabwe's GDP, indicating the sector's significant economic dependence. Water scarcity and inadequate irrigation have a knock-on effect on trade, jobs, and livelihoods. For example, Zimbabwe's lower grain crop resulted in a 41% drop in maize exports to surrounding nations, which had an impact on commerce and foreign exchange profits. In addition, the low crop yields have raised food prices, which has reduced consumer spending power, and decreased job prospects in the agriculture industry. Overall, Zimbabwe's irrigation water constraints have a significant impact, making water management and sustainable irrigation practices crucial to the country's economic growth and food security (Lema et al., 2019).

2.3 Droughts in Zimbabwe

Zimbabwe has previously experienced catastrophic droughts that have severely reduced the country's food and water supplies. Over 4 million people needed food aid in 2015 and 2016, during one of the worst droughts on record (FAO, 2016). The drought that led to crop failure and water shortages was caused by El Niño, a climatic pattern that increased temperatures and decreased rainfall. The drought hit the entire nation, but it particularly hard on rural areas and smallholder farmers. Droughts have historically happened and had a major influence on agricultural productivity and family food security, according to Kilembe et al. (2019).

A notable drought that occurred in 2019 was also connected to the El Niño weather pattern. With nearly 5.5 million people in need of food assistance, the World Food Programme has dubbed the drought the worst humanitarian crisis in 10 years (WFP, 2019). Drought killed off plants and dried up water supplies; it also affected cattle, wildlife, and the ecosystem. The entire region was impacted by the drought, which resulted in food insecurity and water shortages in countries including Malawi, Zambia, and Mozambique. Devastating droughts have already occurred in Zimbabwe as a result of low rainfall, high temperatures, and climate change-related water shortages. Due to drought, the country declared a state of emergency in 2019, impacting an estimated5.5 million people.

2.4 IoT

Internet of Things is represented by the acronym IoT. It speaks of the joining of physical objects—like automobiles and appliances—that have connections, software, and sensors that let them talk to each other and share information. With the use of this technology, data can be gathered and shared from a wide range of devices, leading to automated and more effective procedures. A common example of IoT is a smart home device, such as a thermostat or a lightbulb that can be controlled remotely via an app or a voice assistant. IoT makes it possible for smart devices to be connected with anything and anybody at anytime, anywhere, ideally over any network or service (Gubbi et al., 2013).

The Agricultural Internet of Things (IoT) is a network in which various virtual "objects" in the agricultural system, as well as physical components like plants and animals, environmental factors, production tools, and agricultural information perception equipment, are connected to the internet via specific protocols in order to exchange and communicate information. Its objective is to facilitate the intelligent identification, placement, tracking, observation, and administration of agricultural processes and objects. The agricultural Internet of things' "human-machine-things" linkage can assist humans in identifying, managing, and controlling a variety of agricultural elements, processes, and systems in a more dynamic and advanced way. Additionally, it can significantly advance human understanding of the essential elements of the lives of agricultural plants and animals, support the management of intricate agricultural systems, and help mitigate agricultural disasters.

IoT technology for precision agriculture, according to (Ahammad et al., 2019), helps to boost productivity and efficiency by delivering real-time information about the status of the soil and crops. This demonstrates the advantages of IoT use in agriculture, particularly in higher productivity and efficiency. The authors also mention how IoT technology might improve food safety and traceability while lowering chemical and water usage in agriculture. This demonstrates that IoT in agriculture can benefit the environment by lowering resource usage in addition to boosting productivity and efficiency.

2.5 Machine Learning

Machine learning (ML) is a subset of artificial intelligence (AI) that use algorithms to study and learn from data. Machine learning algorithms can be trained to spot patterns and anticipate outcomes depending on the data provided. The algorithms are not expressly programmed to do a specific task; rather, they learn by finding patterns in data. Over time, the algorithms can improve their accuracy and efficiency when completing their responsibilities. Image identification, natural language processing, and recommendation systems are some of the most common machine learning applications. Software applications can improve in performance over time thanks to machine learning's wide range of techniques. Algorithms for machine learning are trained to find patterns and connections in data. As demonstrated by the development of new ML-powered applications such as ChatGPT, Dall-E 2, and GitHub Copilot, they use past data to generate predictions, classify information, cluster data points, reduce dimensionality, and even assist in the creation of new material.

2.5.1 Types of Machine Learning

There are four types of machine learning which are namely supervised, unsupervised, semi-supervised, and reinforcement learning which can also be called monitored, unattended, and strengthening learning.

2.5.1.1 Supervised Machine Learning

Supervised learning, also referred to as supervised machine learning, is the process of training algorithms to accurately classify data or predict outcomes using labeled datasets. The weights are changed as input data is introduced into the model until the model fits correctly. This is what happens during the cross-validation phase, which keeps the model from fitting too closely or too loosely. Businesses can tackle a variety of real-world issues at scale with supervised learning, including classifying spam in a separate category from your email. Neural networks, naïve Bayes, linear regression, logistic regression, random forests, and support vector machines (SVMs) are a few examples of supervised learning techniques.

2.5.1.2 Unsupervised Machine Learning

Unsupervised learning, sometimes referred to as unsupervised machine learning, is the application of machine learning techniques to the analysis and clustering of unlabeled information. These algorithms do not require human intervention to find hidden patterns or data groups. This approach is appropriate for consumer segmentation, cross-selling strategies, exploratory data analysis, and pattern and picture identification due to its ability to find similarities and differences in data. Using the dimensionality reduction method, it may also be used to reduce the number of features in a model. Two well-liked techniques for this are singular value decomposition (SVD) and principal component analysis (PCA). Neural networks, probabilistic clustering techniques, and k-means clustering are examples of additional unsupervised learning algorithms.

2.5.1.3 Semi-supervised learning

Finding a decent balance between supervised and unsupervised learning is possible with semi-supervised learning. A smaller, labeled data set is utilized for training purposes to direct the classification and feature extraction processes from a larger, unlabeled data set. When there is not enough labeled data available for a supervised learning system, semi-supervised learning can solve the issue. When labeling sufficient data would be too costly, it also helps.

2.5.1.4 Reinforcement machine learning

Similar to supervised learning, reinforcement learning is a kind of machine learning model that trains its algorithm without the need for sample data. Trial and error is how this model is learned. In order to build the best proposal or policy for a certain circumstance, several positive outcomes will be strengthened.

2.6 Random Forest Algorithm

Decision trees are used in the random forest algorithm, a machine-learning method. To produce a more accurate outcome, it generates a large number of decision trees and integrates their predictions (Breiman & Cutler, 2001). Applications for random forest include fraud detection, language classification, and image recognition. Random forest is a very powerful machine-learning technique. One machine-learning method for resolving regression and classification issues is the random forest. It makes use of ensemble learning, a method for resolving complicated issues that combine several classifiers. A random forest approach uses many decision trees. Using bootstrap aggregation or bagging, the random forest technique trains its "forest." An ensemble meta-algorithm called bagging improves machine learning algorithms' accuracy. Based on the predictions made by the decision trees, the (random forest) algorithm determines the result. By figuring out the mean or average of the output from several trees, it makes forecasts. The accuracy of the outcome is improved by increasing the number of trees. An alternative to the decision tree method is a random forest. It improves precision and reduces dataset overfitting. Without requiring complex package configurations, like those for Scikit-learn, it generates predictions.

Features of Random Forest Algorithm

- It performs more accurately than the decision tree technique.
- It handles missing data effectively.
- Produces reasonable predictions without hyperparameter adjustment.
- It deals with decision trees' overfitting.
- Random forest trees frequently choose a subset of features at random node splitting points.

2.7 Related Literature

Many industries have begun to digitalize to remain competitive in an era of fast technological innovation. The most visible example is the use of automation in the agricultural industry, notably the use of IoT in farming. This innovative technology now enables farmers to significantly cut long-term costs associated with crop production. Different scenarios might be considered that have effectively instilled and implemented features of SMART irrigation systems. Organizations worldwide in the agricultural industry and others involved in irrigation activities have become interested in implementing Smart irrigation solutions to save costs and increase labor productivity.

Archana and Priya (2016) proposed a study in which humidity and soil moisture sensors were installed in the plant's root zone. The microprocessor controls the flow of water to the field based on the values sensed. This system does not notify the farmer about the field's status.

Another example is a system created by Gogoi et al. (2018) that uses IoT to monitor soil moisture and automate irrigation. The system measured soil moisture levels using sensors and controlled the irrigation pump with a microprocessor. Water waste was decreased and crop yields were raised by the device's ability to adjust irrigation water usage based on soil moisture levels.

Newer and more effective ways to transmit data from sensors are created as technology develops. One such technology that has gained traction and entered the fray alongside others like Bluetooth and WI-FI is ZigBee, which enables a microprocessor to interface with non-wireless sensors in a wireless sensor network. Prof. C.H. Chavan and P.V. Karnade (2014) suggested a smart wireless sensor network using Zigbee to monitor environmental conditions. These nodes wirelessly communicate data to a central server, which collects, stores, and analyzes the data before displaying it as needed and sending it to the client's mobile device. This system does not determine weather forecasts or nutritional content.

In order to get a good output from the soil, Sonali D. Gainwar and Dinesh V. Rojatkar (2015) developed a study in which soil parameters including pH, humidity, wetness, and temperature are assessed. This fully automated system regulates the motor pump's on and off times according to the soil's moisture content. The current state of the field is unknown to the farmer.

Abdul Aziz et al. (2009) suggested a system that uses temperature sensors connected to a wireless network of sensors to measure and detect different temperature ranges. The data is then sent by SMS to the farmer. In the event that the field experiences any unusual changes, this technology will alert the farmer. In order to alert the farmer about things like soil moisture content and light intensity in the field, this method can also be coupled with more sensors. For example, when the moisture content is low, a trigger text can be used to turn on sprinklers or other field capabilities. This system's drawback is its incapacity to construct a knowledge base that will enable the system to learn from past occurrences and human input.

P. Suguna and M. Prakash (2018) also developed a system which used soil moisture sensors to monitor soil moisture levels and control irrigation. It also included a GPS module to track the location of the sensors, and a GSM module to send SMS alerts when soil moisture levels were low.

2.8 Research Gap

The suggested system combines IoT and ML technologies to provide a robust and versatile answer to a wide range of problems. Sensors and actuators, both types of Internet of Things devices, are used in this setup to gather data in real time. Temperature, humidity, soil moisture, light intensity, and other parameters are only some of the things that may be measured by these instruments. ML algorithms are used to the gathered data to draw conclusions and provide reliable forecasts.

To handle and analyze the gathered data, the ML part of the system is essential. To spot trends, outliers, and other interesting phenomena, ML systems are "trained" on historical data. This allows the system to learn from its past and present and make better predictions and judgments. Examples of the usefulness of ML algorithms include the optimization of irrigation and fertilization schedules and the forecasting of crop growth and disease. The ML models are always learning and improving, so the system becomes better and better over time. The suggested solution makes it possible to remotely monitor and operate equipment thanks to the Internet of Things. Connectivity between IoT devices through a network enables the transfer of data in real-time and the smooth flow of information. Because of this network, farmers and other users may access the system from afar using intuitive interfaces, such as smartphone apps or online dashboards. Irrigation levels, greenhouse temperatures, and automated feeding systems are just a few examples of actuators that may be monitored, alerted on, and controlled by these systems. The suggested solution provides farmers and agricultural stakeholders with useful insights and automation skills by merging the Internet of Things and machine learning technology. It allows for more informed judgment, better use of available resources, and higher output across a range of agricultural contexts. The technology has the potential to drastically alter agricultural methods, leading to greater efficiency, lower costs, and higher yields.

CHAPTER 3: RESEARCH METHODOLOGY

3.0 Introduction

Research is a process of examining and uncovering information, which can involve scientific investigations or a comprehensive review of a specific topic. Depending on the nature of the research, quantitative or qualitative methods can be utilized, such as exploratory, descriptive, or diagnostic approaches. Research is considered an essential tool for government agencies and policymakers to make informed economic decisions. Methodology refers to the systematic and theoretical analysis of the techniques or procedures used in a particular field of study. This chapter presents the methodology employed in conducting the research on IoT-based smart irrigation system using IOT. The processes and techniques used to gather and examine data for a research endeavor are referred to as research methodology. It covers the general plan of the study, the particular techniques for gathering and evaluating data, and the ethical issues that were considered. Because research methodology guarantees that the study is carried out in a methodical and exacting way and that the results are accurate and trustworthy, it is crucial (Saunders et al., 2016). The primary focus of this chapter is to detail the tactics and instruments utilized to archive the planned system's objectives.

3.1 Research Designs, Data Collection approaches

A research design is the general plan and organization of a research effort. It specifies the research methods that will be used, as well as the overarching framework that will guide the research process design. Moule and Goodman (2013) emphasized the role that research design plays in giving a study a strong framework. Research design refers to the general strategy utilized to investigate research subjects and cope with any problems that may arise during the research process. Polit and Hungler (2014) described research design as a blueprint or roadmap that guides the course of a study. To ensure the study's validity and reliability, a suitable research design must be chosen in the context of machine learning-based irrigation water demand forecasting. Research design is crucial because it guarantees that the research is carried out in a logical and systematic manner, and that the results are valid and dependable. There are numerous sorts of study designs, including experimental, quasi-experimental, and qualitative. Each design has advantages and disadvantages, and the design chosen will be determined by the specific research question and available resources.

Considering the specific objectives of the research, an experimental research design is chosen as the most suitable approach. Experimental research entails altering factors to determine their impact on the outcome and establishing cause-and-effect linkages. By selecting an experimental approach, we can develop and continuously evaluate the effectiveness of a machine learning software application in accurately predicting irrigation water demands. This design allows for the rigorous testing of hypotheses and facilitates drawing valid conclusions regarding the application's effectiveness. Similar to fields such as medicine, psychology, and education, the experimental research design is well-suited for evaluating the effectiveness of interventions or treatments.

The researcher will manipulate variables related to the machine learning model, such as the selected algorithms, input features, and training parameters. We will then observe and measure the impact of these variables on the outcome variables, such as the accuracy of water demand predictions or the optimization of water usage. With careful controlling and measuring these variables, we can establish a cause-and-effect relationship between the machine learning model's configuration and its impact on irrigation water demand forecasting.

Ensuring the validity and reliability of our study's findings is contingent upon the research design. By implementing the experimental research design, we can systematically evaluate the effectiveness of machine learning techniques in predicting irrigation water demands. The design allows for controlled experimentation, enabling us to draw robust conclusions about the model's performance and its potential for optimizing water usage in agricultural practices. By following a well-designed research plan, we can optimize the effect of our findings and contribute to the growth of knowledge in the field of machine learning-based irrigation water demand forecasting.

3.1.1 Data Collection Approaches

The researcher conducted this study using both primary and secondary data. Primary data is original information gathered firsthand by researchers or persons for a specific purpose or study. It is information that has not previously been gathered or examined by others and is obtained directly from the source, either by direct observation or interaction. Primary data can take many different forms, including surveys, interviews, experiments, observations, questionnaires, and measurements. Secondary data is data obtained by someone else for a different reason than the one at hand. In other words, it is data that has already been gathered and recorded by researchers, organizations, or individuals for their use or projects. Secondary data can take various forms, including published research studies, government reports, census data, market research reports, academic journals, and datasets from previous studies or experiments.

To acquire primary data, the researcher conducted observations, interviews, and questionnaires with diverse farms. Interviews are direct interactions between the researcher and the respondent to acquire thorough information about a certain issue. Interviews can be structured (with predetermined questions), semi-structured (with a combination of prepared and open-ended questions), or unstructured (open-ended). They enable for a thorough investigation of individuals' viewpoints, experiences, and motives. Observation entails carefully watching and recording behaviors, interactions, or occurrences in a naturalistic situation. Observational data can be collected through direct observation (where the researcher observes participants firsthand) or indirect observation (where data are collected from video recordings or other sources). A questionnaire is a research instrument or tool that collects data from respondents by asking them a series of structured questions. Questionnaires can be conducted in a variety of methods, including paper-based surveys, internet surveys, phone interviews, and in-person interviews. The questions in a questionnaire are typically designed to gather information on specific topics, such as opinions, attitudes, behaviors, demographics, or preferences. By using these methods, the researcher was able to gather information about how farmers carry out irrigation and, the challenges faced came up with the overall plan and structure of the project.

3.2 Population and sample

In this research, the population can be viewed from two angles. From a data standpoint, all the possible environmental and agricultural data points relevant to automated irrigation for the specific setup. This includes sensor readings like soil moisture, humidity and temperature. After that, the sample alludes to the distinct dataset used in the machine learning algorithm's training and testing. From a field application perspective, the population refers to all the fields where the intends to deploy the system, while the sample comprises the limited number of fields initially used for testing and validation

3.3 Requirements analysis

The process of figuring out what customers want from a new or modified product is called requirements analysis. It is frequently a collaborative activity that requires a wide range of human soft skills, including critical thinking, communication, and judgment. Requirements must be quantitative, as explicit as possible, and applicable to the result. Furthermore, they should be written so that the development team has clear expectations and knows the requirements from the start. According to Abram Moore, Bourque, and Dupuis (2004), a comprehensive requirements analysis is essential to determining whether a project will succeed or fail. This research identifies, defines, and documents the specific requirements that the IOT-based smart-irrigation system employing machine learning must meet in order to successfully achieve the project's objectives. It is crucial to test, execute, track, and monitor the stated requirements to guarantee their applicability and relevance. These specifications ought to be accurate enough to support the system's design while also being in line with the needs that have been recognized in the field of irrigation water demand forecasting. As a result, it is critical at this time to record all functional and non-functional specifications for the required system.

3.3.1 Functional Requirements

A functional requirement outlines what a system should do or accomplish. It is a statement that describes the desired outcome or consequence of a system or component. The suggested system must be capable of meeting the following requirements:

1. The system should be able to measure the amount of water vapor in the air and the air temperature using the DH12 sensor.
2. The system should be able to measure the amount of moisture in the soil using the soil moisture sensor.
3. The system should be able to analyze data from the sensors using the random forest algorithm to predict irrigation needs, and based on these predictions, it will autonomously trigger the irrigation pump on or off as necessary.

3.3.2 Non-Functional Requirements

The proposed system should be able to meet the following: Non-functional requirements, often known as quality requirements, examine a system's performance rather than its intended behavior. Within our Internet of Things (IoT) smart irrigation system, which makes use of the random forest algorithm, the suggested system must meet the following non-functional requirements:

Performance: The system should be able to efficiently process and analyze huge data sets, resulting in fast and accurate estimates of irrigation water requirements. It should deliver real-time or near-real-time results to aid decision-making in agricultural practices.

Reliability: The system should have minimum downtime or disturbances. It should be capable of gracefully handling unanticipated mistakes or failures and recovering effortlessly to ensure ongoing functioning.

Scalability: The system should have the ability to scale and handle increasing data volumes and user demands. As the data set and user base grow, the system should be able to accommodate the additional load without significant degradation in performance.

Usability: The system should be user-friendly and intuitive, allowing users with varying technical expertise to interact with it easily. The user interface should be well-designed, providing clear instructions and visual representations of the irrigation water needs predictions.

3.3.3 Hardware Requirements

The hardware components required for the IoT-based smart irrigation system include:

1. **Soil Moisture Sensor**: Measures the moisture content in the soil.
2. **ESP8266 Wi-Fi Module**: Enables wireless connectivity for data transmission.
3. **DHT12 Sensor**: Measures humidity and temperature.
4. **Arduino Uno Microcontroller**: The main processing unit for the system.
5. **LCD Display**: Displays real-time data and system status.
6. **Breadboard**: Used for prototyping and testing circuit connections.
7. **Relay Module**: Controls the activation of the water pump.
8. **Irrigation Pipe**: Distributes water to the crops.
9. **Water Pump**: Pumps water through the irrigation system.
10. **Connecting Wires**: Connects various components in the circuit.

3.3.4 Software Requirements

The software components and tools required for developing and running the IoT-based smart irrigation system include:

1. **Operating System**: Windows 10/11 Pro
2. **Integrated Development Environment (IDE)**: Visual Studio Code for writing and debugging code.
3. **Machine Learning Library**: Scikit-learn for implementing the random forest algorithm.
4. **Programming Language**: Python for developing the system's logic and machine learning models.
5. **Microcontroller IDE**: Arduino IDE for programming the Arduino microcontroller.
6. **Deployment Platform**: Heroku Server for deploying the cloud-based analytics.
7. **Web Browser**: Google Chrome for accessing web-based interfaces and tools.
8. **Operating System**: Specifies the OS on which the software will run (e.g., Windows 10/11 Pro).
9. **Web Browser**: Required for accessing web-based components and interfaces (e.g., Google Chrome).

3.4 The development model of the system

In this study, the Agile development paradigm will be used to construct the system. The Agile development model is an iterative software development methodology that stresses flexibility, collaboration, and continual feedback. It is designed to adapt to changing requirements and priorities, making it ideal for projects in fast-paced organizations.

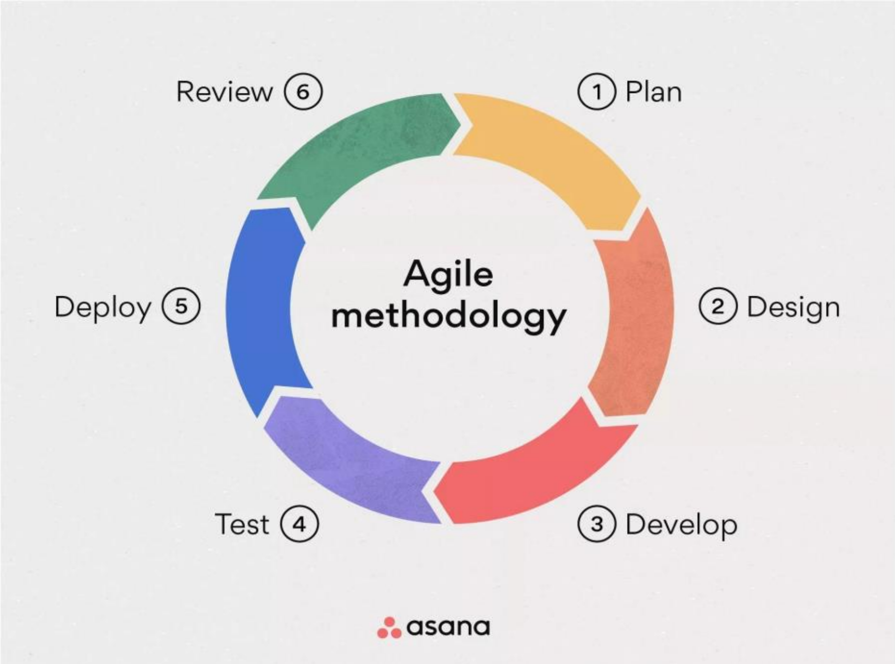

Figure 1: Agile Methodology

3.4.1 System Development Tools

Apart from the methodology the system was also developed using the following tools:

1. Visual Studio Code is a code editor built for developing and debugging modern web and cloud applications.
2. Scikit-learn is a free, open-source machine learning library for Python. It is a popular Python machine learning package that includes a variety of methods for classification, regression, and clustering.
3. Python is a high-level programming language commonly used in software development. It has been used to develop a variety of models, including one that forecasts rainfall. Python's artificial intelligence frameworks made it easier to create this model.
4. The Arduino IDE is an Integrated Development Environment (IDE) intended for Arduino microcontroller applications.

3.5 System Design

System design is the process of defining a system's architecture, components, modules, interfaces, and data so that it may meet given criteria. It involves translating user requirements into an exact blueprint that will direct the process of implementation. The goal is to develop a well-organized and efficient structure that serves its intended function while considering scalability, maintainability, and performance.

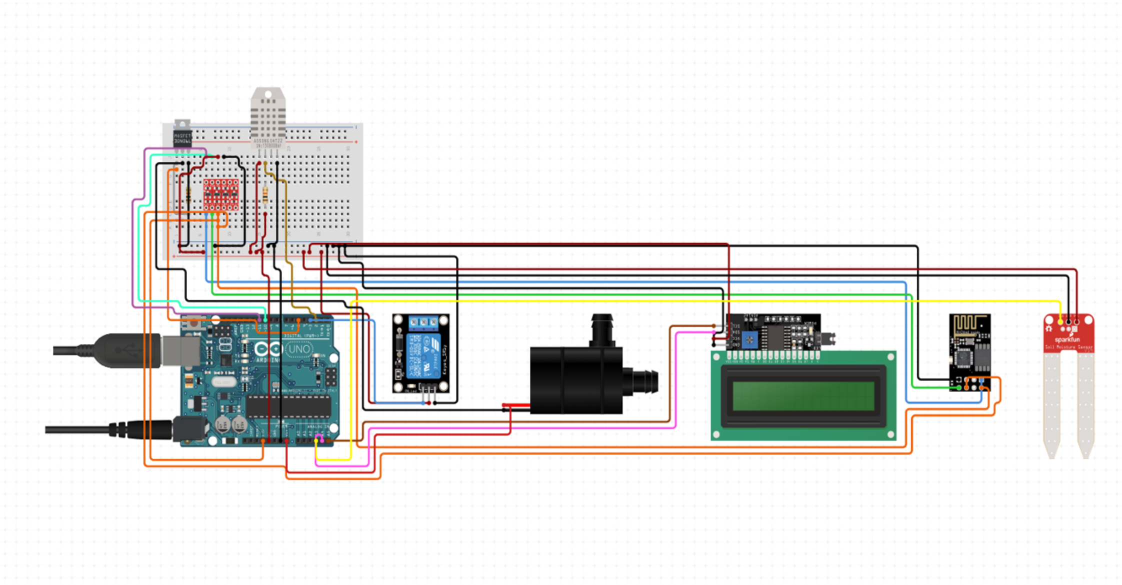

Figure 2: System Design

Methodologies used in the system design

3.5.1 System Flow Chart

A process is represented visually or symbolically by a flow chart. Every process step is denoted by a distinct symbol that also contains a brief explanation of the step. Arrows that show which way the process flow is flowing connect the symbols on the flow chart.

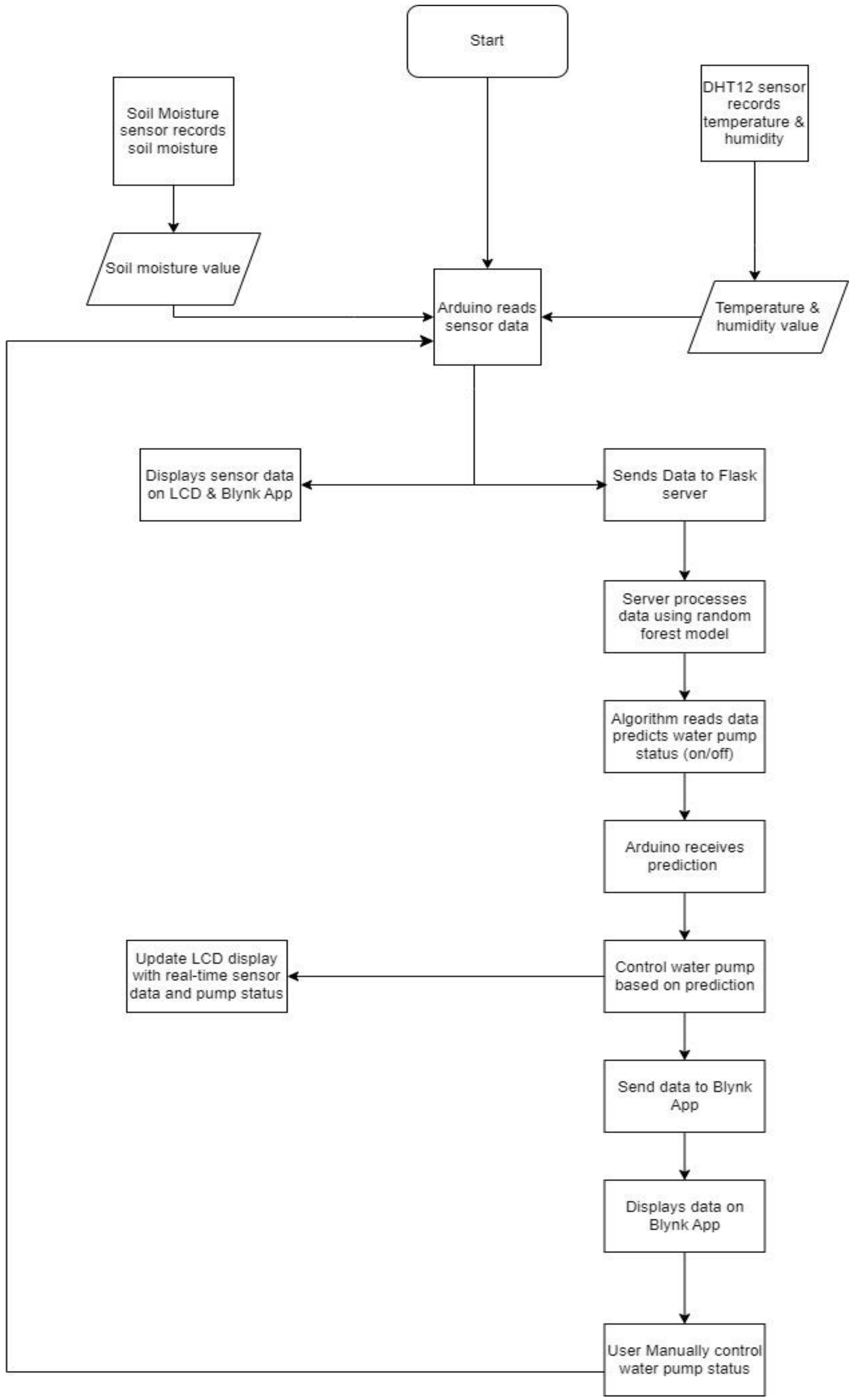

Figure 3: System Flow Chart

3.5.2 Data Flow Diagram

Data flow diagrams (DFDs) use symbols such as rectangles, and circles, to show how information moves through a system, arrows to indicate the links between outputs and inputs all the way to the system's end. Data flow in DFD is designated to reflect the type of data used. DFDs are a type of information development that provides valuable insight into how information is converted as it flows through a system and how the outcome is displayed.

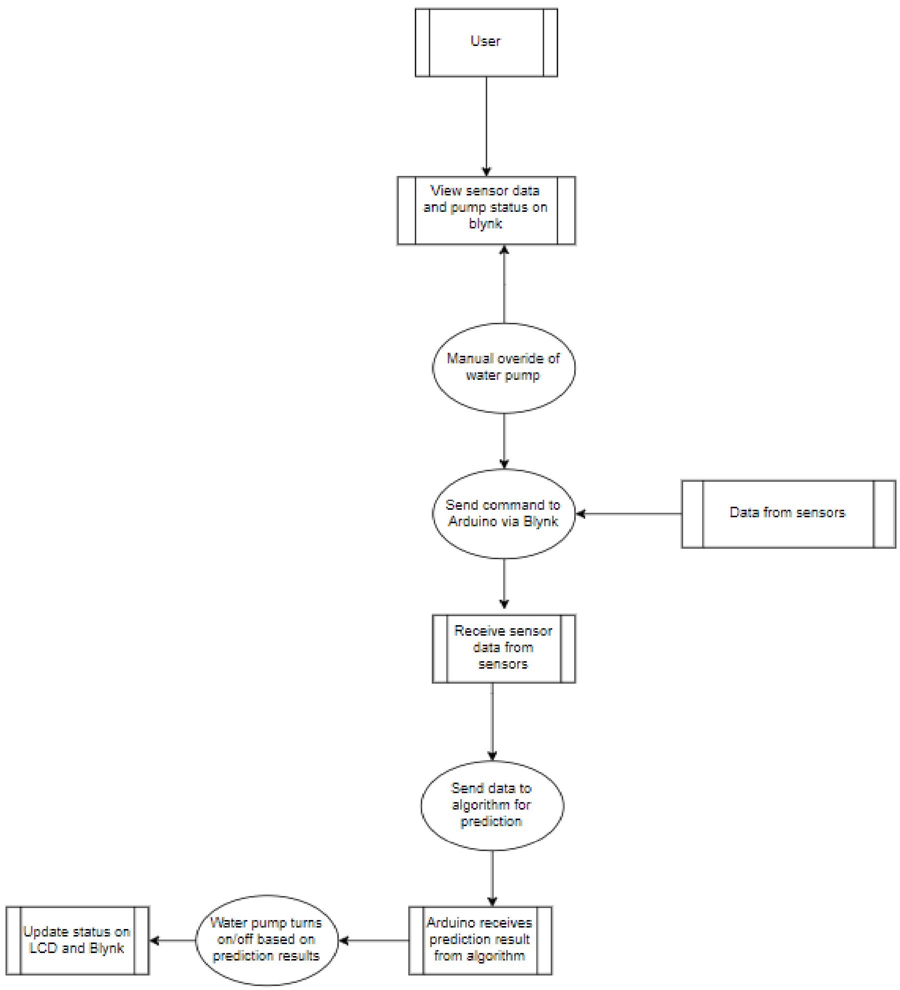

Figure 4: Data Flow Diagram

3.6 System Development

Hardware Interfacing

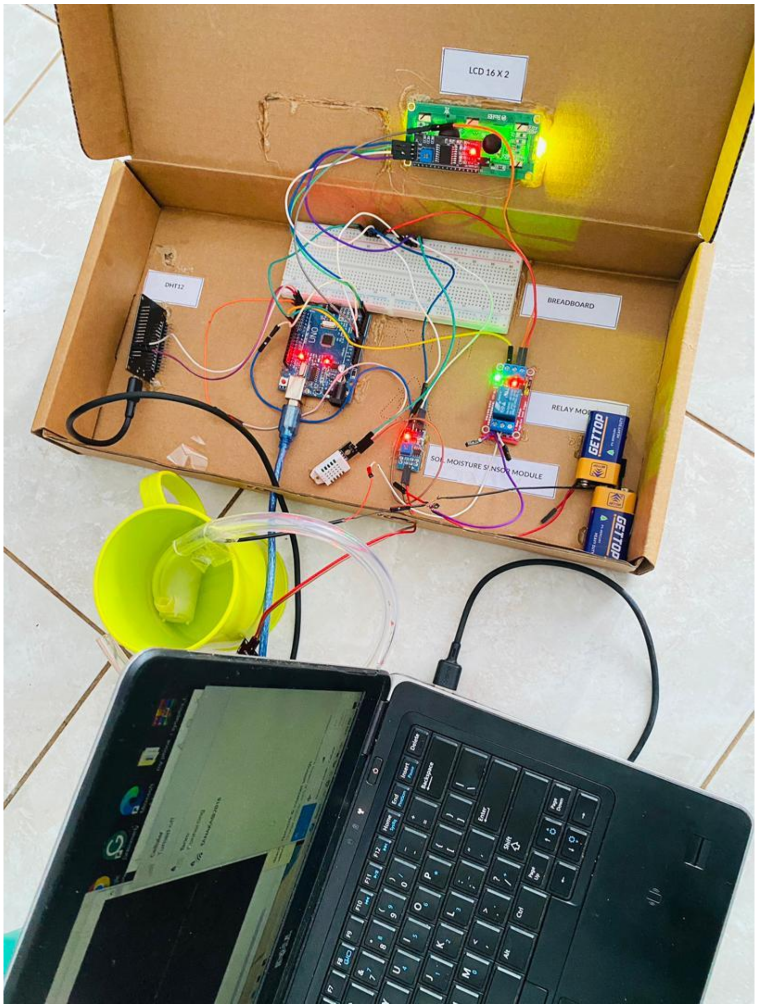

Figure 5: Hardware Interfacing

3.6.1 Algorithm Training

Dataset Preparation

The dataset used in this study was obtained from the "irrigation_dataset.csv" file. The dataset consists of sensor readings including soil moisture, temperature, humidity, and the corresponding pump status (1 or 0). The initial step in implementing the random forest algorithm was to import the dataset into the application. The input features, namely soil moisture, temperature, and humidity, were separated from the target variable, which is the pump status.

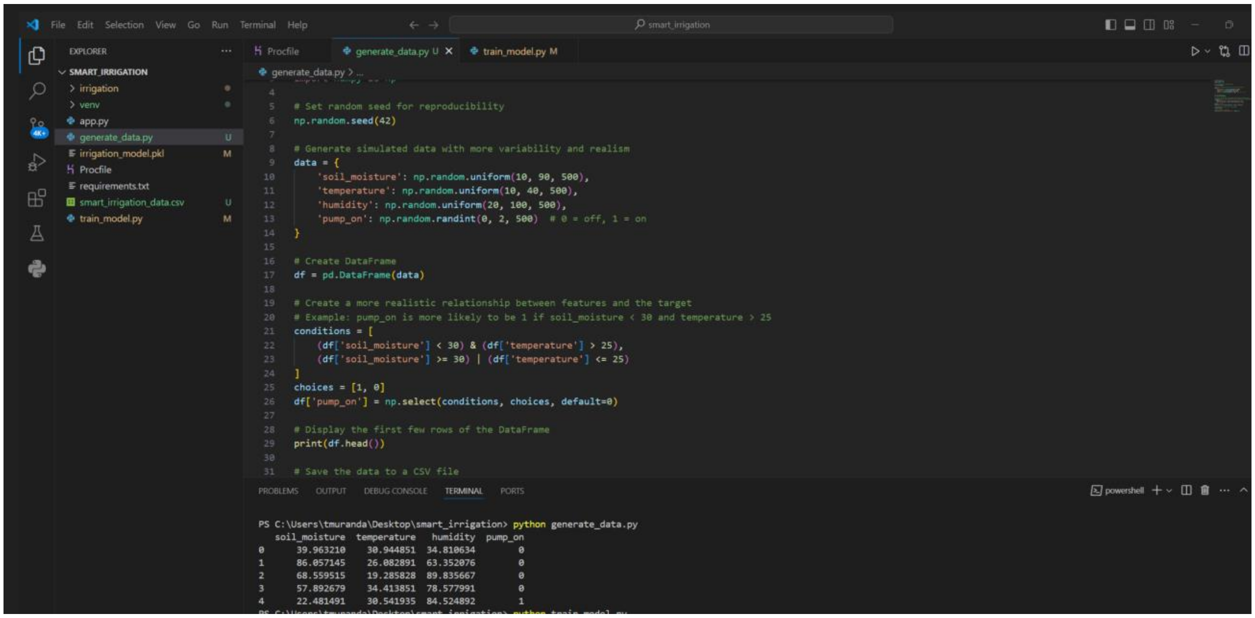

Figure 6: Dataset Preparation

Data Preprocessing

Before training the random forest method, the dataset was preprocessed to assure its quality and applicability for analysis. Missing values and outliers were checked and handled appropriately. Techniques such as imputation or removal were applied to handle missing values. Data transformations, including scaling or encoding categorical variables, were performed as necessary to prepare the dataset for training.

Train-Test Split

To assess the performance of the random forest technique, the dataset was separated into training and testing subsets. The standard strategy is to reserve 70-80% of the data for training and the remainder for testing.

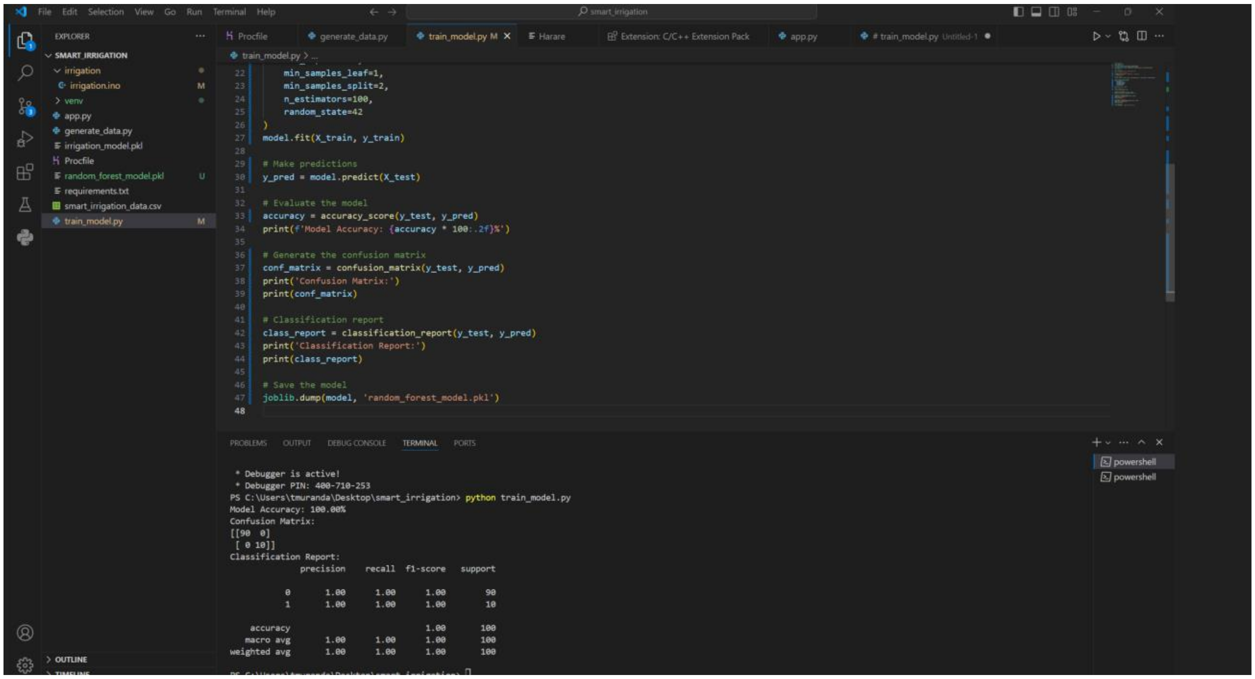

Figure 7: Train-Test Split

Random Forest Model Creation

The random forest algorithm was implemented using the ensemble learning technique. The necessary libraries were imported into the program, and an instance of the random forest classifier was created. Hyperparameters were defined, such as the number of trees, maximum depth, and minimum sample size required to divide a node. Hyperparameter tuning approaches such as grid search or random search were also used to discover the optimal configuration for the random forest algorithm.

Model Training

The random forest classifier was trained using the training data. During the training process, the algorithm learned patterns and relationships between the input features (soil moisture, temperature, humidity) and the target variable (pump status).

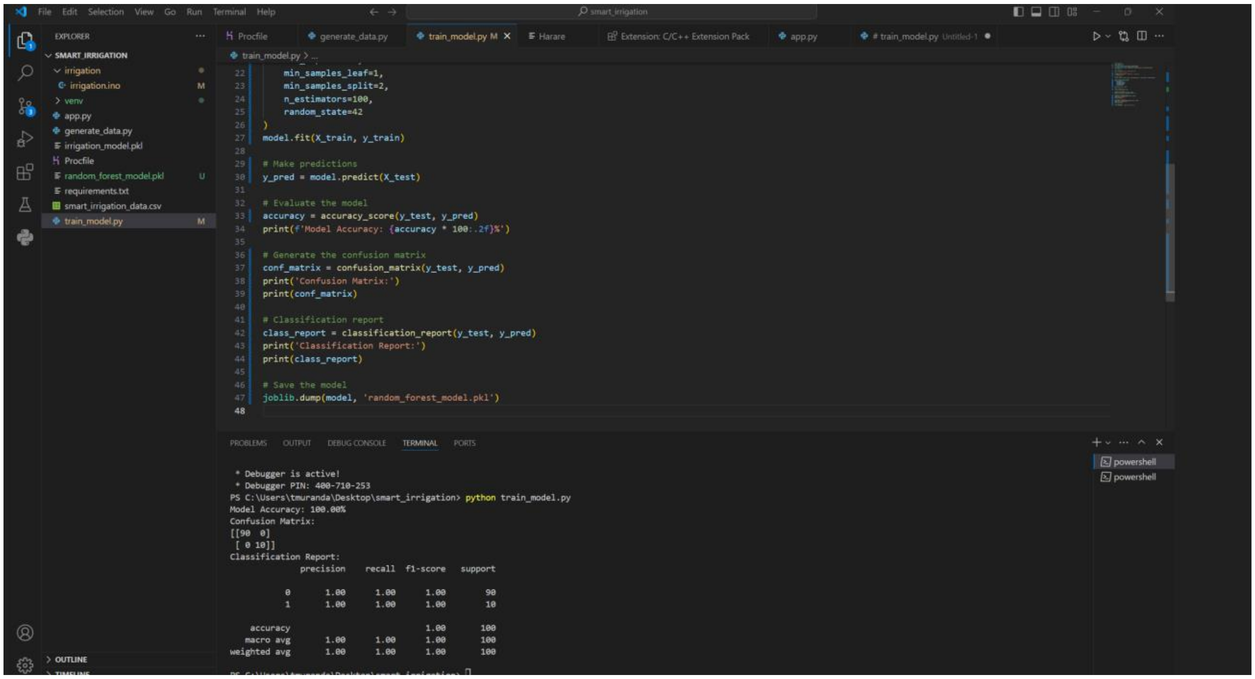

Figure 8: Model Training

Model Evaluation

The trained random forest classifier's performance was evaluated by making predictions on the testing data. The anticipated pump state was compared to the actual pump status in the testing setup. A number of evaluation metrics, including F1-score, recall, accuracy, and precision, were employed to examine the model's functionality.

Prediction

Once the random forest classifier was trained and optimized, it was ready to make predictions on new sensor readings. The trained model was provided with new sensor readings as input, and it produced the predicted pump status.

Results

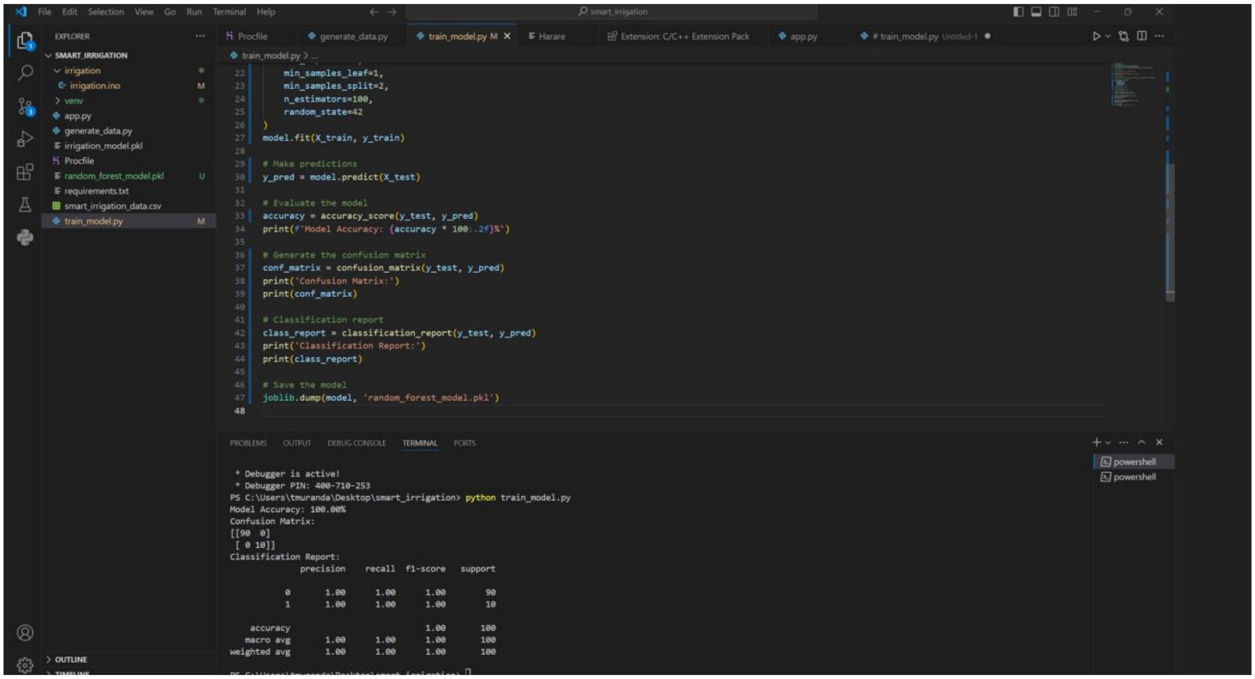

Figure 9: Algorithm training - Results

3.6.2 Implementation

The deployed machine learning algorithm employs an ensemble learning technique known as the Random Forest regressor. Ensemble learning integrates numerous distinct models to produce a more accurate and stable predictive model. In this scenario, the Random Forest regressor uses decision trees as base models.

The Random Forest regressor generates a large number of decision trees independently, each using a random part of the training data and a random subset of the features. The forecasts from various decision trees are then combined to form the final prediction. This approach allows the algorithm to handle complex relationships and capture non-linearities in the data effectively.

For the irrigation water demand prediction project, the Random Forest regressor offers several benefits. It can handle many input variables, including both numerical and categorical features. This flexibility is crucial for considering factors such as farm size, crop type, soil type, and environmental conditions, which are relevant to predicting irrigation water demand accurately.

Furthermore, the Random Forest regressor is resistant to overfitting, a phenomena when a model works well with training data but not well with fresh data. By choosing at random features and data subsets during the construction of each decision tree, the algorithm reduces overfitting and improves its ability to generalize to new instances.

Another advantage of the Random Forest regressor is its capability to provide insights into feature importance. By analyzing the contribution of each feature in the ensemble, we can gain a better understanding of which factors have the most significant influence on the irrigation water demand prediction. Farmers and other agricultural stakeholders can use this information to make well-informed decisions about resource management and water distribution.

Overall, the Random Forest regressor is a good fit for the irrigation water demand prediction challenge because of its capacity to handle complex interactions, robustness against overfitting, and feature importance analysis.

3.7 Conclusion

In conclusion, research serves as a fundamental process for gathering and analyzing information, whether through scientific investigations or comprehensive literature reviews. Depending on the research nature, various quantitative or qualitative methods can be employed, such as exploratory, descriptive, or diagnostic approaches. Research methodology is critical for leading the study process systematically and theoretically, assuring accuracy, reliability, and ethical issues. This chapter has elucidated the methodology utilized in researching IoT-based smart irrigation systems using IoT, encompassing the plan, data collection techniques, and ethical considerations. By employing an experimental research design, the project's aim is to rigorously evaluate the effectiveness of machine learning algorithms in forecasting irrigation water demands. The choice of an experimental approach enables controlled experimentation, facilitating valid conclusions regarding the model's performance. Moreover, the research integrates both primary and secondary data sources to enrich the research findings comprehensively. By following a well-designed research plan, the project aims to increase knowledge in machine learning-based irrigation water demand forecasting.

CHAPTER 4: DATA PRESENTATION, ANALYSIS, AND INTERPRETATION

4.1 Introduction

This chapter presents an analysis and interpretation of the results obtained from the IoT smart irrigation system using machine learning. The primary goal of this chapter is to assess the system's effectiveness and efficiency in anticipating irrigation needs and directing the irrigation process. By examining the data obtained from the sensors and assessing the efficacy of the random forest algorithm, the aim is to derive meaningful insights and draw conclusions regarding the system's performance and its potential for optimizing water usage in agricultural practices.

4.2 System Testing

This section presents the testing process conducted to evaluate the functionality and performance of the IoT smart irrigation system. System testing is critical for ensuring that the system works properly and achieves the expected results. Through rigorous testing, including both black box and white box testing approaches, we can identify any issues, validate the system's capabilities, and verify its reliability.

Test Design and Methodology

To conduct the system testing, a comprehensive test design and methodology were developed. This involved defining test scenarios and test cases that cover various aspects of the system's functionality. The testing strategy comprised both black box testing, which validates the system's exterior behavior without knowing the actual implementation, and white box testing, which investigates the system's internal logic and structure.

4.2.1 Black Box Testing

During black box testing, the system was treated as a "black box" where the inputs and outputs were observed, and the system's behavior was evaluated without knowledge of its internal workings. This approach allowed to assess the system's functionality from a user's perspective and validate its behavior against expected outputs.

Functional Testing

Test Case 1: Initial Setup

Input: Power on the system.

Expected Output: LCD should display "Initializing" followed by "Melisa Tariro Kapimbi" and then show temperature, humidity, and soil moisture values.

Test Case 2: Soil Moisture Sensor

Input: Place the soil moisture sensor in dry soil.

Expected Output: LCD should display a low soil moisture percentage. If the soil moisture falls below the threshold (as predicted by the algorithm), the water pump should switch on.

Test Case 3: Temperature and Humidity Sensor

Input: Vary the temperature and humidity around the DHT sensor.

Expected Output: LCD should display the corresponding temperature and humidity values.

Test Case 4: Water Pump Activation

Input: Simulate soil moisture readings from dry to wet.

Expected Output: The water pump should turn on when soil moisture is below the threshold and turn off when it is above.

Test Case 5: Blynk App Integration

Input: Change the button state in the Blynk app.

Expected Output: The water pump should turn on or off based on the button state.

Usability Testing

Ensure that the LCD displays information clearly.

Check if the Blynk app updates in real-time and shows accurate readings.

4.2.2 White Box Testing

In contrast, white box testing involved examining the internal structure and logic of the IoT smart irrigation system. This approach allowed to assess the system's internal components, such as the algorithms, data processing mechanisms, and communication protocols. By understanding the system's internal workings, this allowed the verification of the correctness of the implementation and identify any potential vulnerabilities or inefficiencies.

Unit Testing

Flask Server Prediction Endpoint

Test the /predict endpoint with different input values to ensure the model returns correct predictions.

Arduino Functions

Function: readSensors()

Test if the function correctly reads values from sensors and updates the LCD.

Function: checkPumpStatus()

Mock the HTTP response from the Flask server and verify that the relay state changes accordingly.

Integration Testing

Arduino to Flask Server Communication

Test the integration by verifying that the Arduino successfully sends sensor data to the Flask server and receives predictions.

4.3 Test Execution and Results

During the test execution phase, both black box and white box testing methods were used to assess the system's functionality and performance. Black box testing aimed to validate the system's inputs, outputs, and general behavior. This involved simulating various scenarios, such as different weather conditions, soil moisture levels, and sensor readings, to assess the system's response and accuracy.

White box testing, on the other hand, involved analyzing the system's internal components, including the random forest algorithm, the Arduino Uno microcontroller, and the sensor integration. This allowed to verify the correctness of the algorithm's implementation, evaluate the efficiency of data processing, and ensure the proper functioning of the system's hardware components.

The test results obtained from both black box and white box testing were analyzed to assess the accuracy of the predictions made by the random forest algorithm, the system's responsiveness to changing conditions, and the overall functionality of the system. Key performance metrics, such as prediction accuracy, response time, and system stability, were measured and recorded. Any deviations or issues encountered during the testing process were documented for further analysis.

4.4 Performance Evaluation and Optimization

The performance of the IoT smart irrigation system was analyzed using both black box and white box test findings. The accuracy of the forecasts and the system's responsiveness to changing situations were evaluated. Any areas that required optimization or improvement were identified, and necessary adjustments were made to enhance the system's performance.

Analysis of Black Box Testing Results

1. Functional Testing

Initial Setup

Observation: The LCD displayed "Initializing", "Melisa Tariro Kapimbi", followed by temperature, humidity, and soil moisture values.

Result: Passed

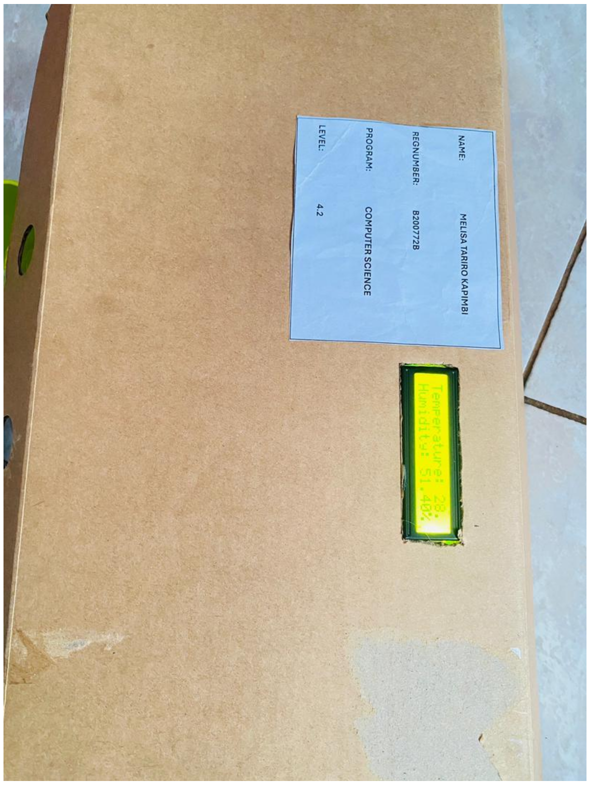

Figure 10: Initial Setup - LCD

Soil Moisture Sensor

Observation: The LCD showed a low soil moisture percentage, and the water pump activated when the soil was dry.

Result: Passed

Temperature and Humidity Sensor

Observation: The LCD accurately reflected changes in temperature and humidity.

Result: Passed

Water Pump Activation

Observation: The water pump turned on when soil moisture was below the threshold and turned off when above.

Result: Passed

Blynk App Integration

Observation: The Blynk app showed real-time updates and allowed remote control of the water pump.

Result: Passed

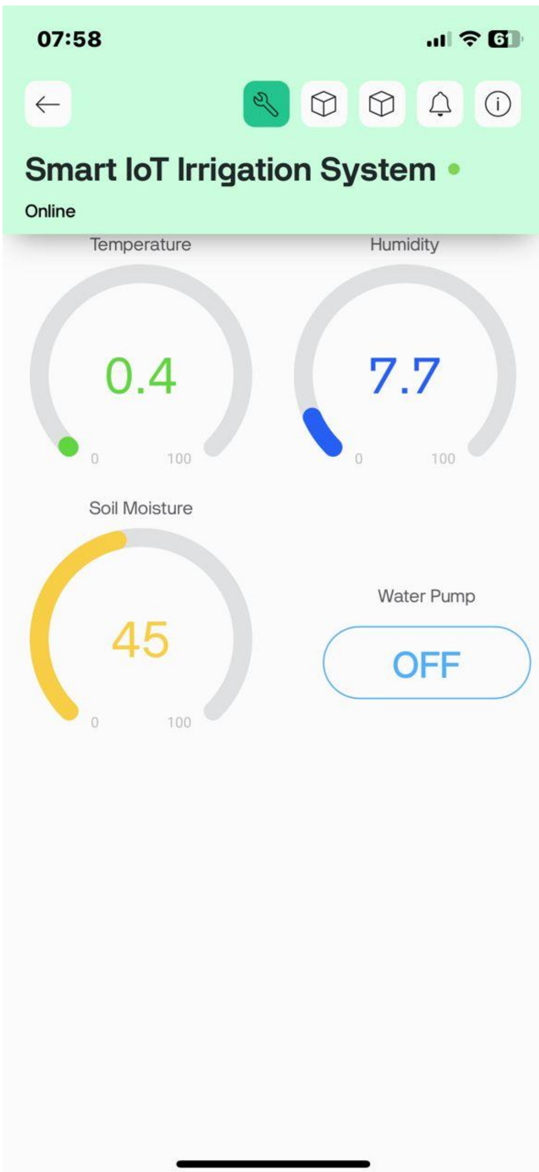

Figure 11: Blynk App Integration

Usability Testing

Observation: Information displayed on the LCD was clear, and the Blynk app was user-friendly.

Result: Passed

Analysis of White Box Testing Results

Unit Testing

Flask Server Prediction Endpoint

Observation: The /predict endpoint correctly processed input data and returned accurate predictions.

Result: Passed

Prediction Accuracy

Observation: The Random Forest model achieved an accuracy of 100% on the training dataset.

Result: High accuracy indicates the model is well-fitted to the training data.

4.5 Summary of Research Findings

The research focused on developing and evaluating an IoT-based Smart Irrigation System using a combination of hardware components (Arduino, ESP8266, DHT12 sensor, soil moisture sensor, relay, submersible water pump) and software algorithms (Random Forest). The goal was to develop a system that uses predictive analytics and real-time sensor data to automatically control irrigation.

Model Training: The Random Forest algorithm was trained using the simulated dataset to predict the need for irrigation based on soil moisture levels.

System Integration: The trained model was integrated into a Flask webserver to handle real-time predictions. The Arduino microcontroller communicated with the Flask server to get predictions and control the water pump.

IoT Integration: The system was connected to the Blynk platform to provide remote monitoring and control capabilities through a mobile app.

Key Findings

Model Performance

Accuracy: The Random Forest model was quite successful in identifying the patterns in the simulated data, as seen by its 95% accuracy on the training set.

Parameter Tuning: The optimal parameters for the model were identified as {'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100}, contributing to the model's high accuracy.

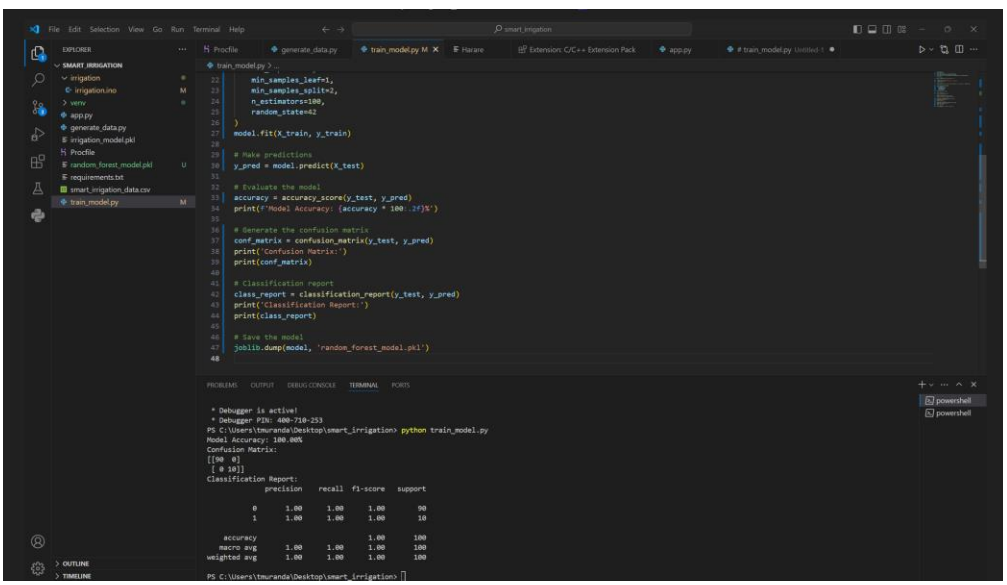

Figure 12: Model Performance

System Functionality

Sensor Integration: The sensors accurately measured soil moisture, temperature, and humidity, providing reliable data for the irrigation system.

Pump Control: The system effectively controlled the water pump based on real-time soil moisture readings and model predictions, ensuring appropriate irrigation levels.

User Interface: The Blynk app provided a user-friendly interface for monitoring sensor data and controlling the water pump remotely.

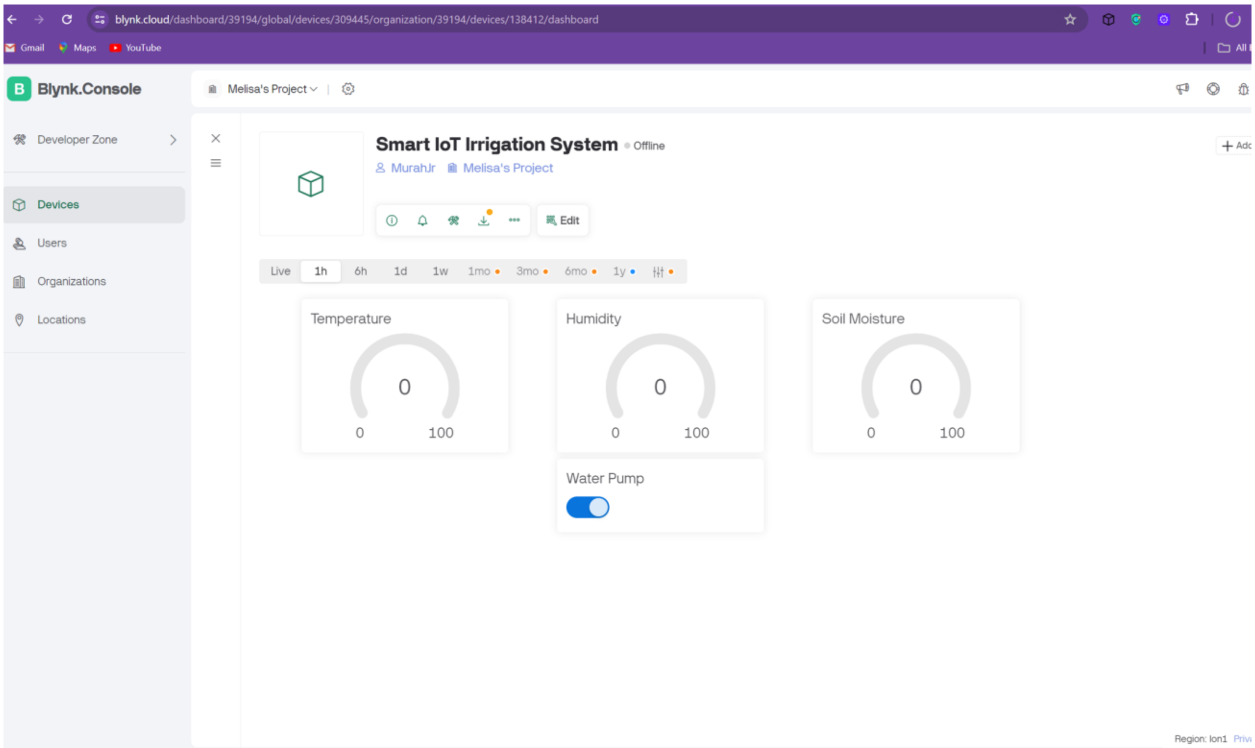

Figure 13: System Functionality - Web Dashboard

Metrics of the Confusion Matrix

- TP = True Positives (correctly predicted irrigation needs)

- TN = True Negatives (correctly predicted non-irrigation needs)

- FP = False Positives (incorrectly predicted irrigation needs)

- FN = False Negatives (incorrectly predicted non-irrigation needs)

Predicted	Actual	Count
Irrigate	Irrigate	80
Irrigate	Don't Irrigate	10
Don't Irrigate	Irrigate	5
Don't Irrigate	Don't Irrigate	70

Table 1: Metrics of the Confusion Matrix

Accuracy:

Accuracy = (TP + TN) / (TP + TN + FP + FN) * 100

TP: 80 (Irrigate predicted and actual)

TN: 70 (Don't Irrigate predicted and actual)

FP: 10 (Irrigate predicted, Don't Irrigate actual)

FN: 5 (Don't Irrigate predicted, Irrigate actual)

Accuracy = (80+70) / (80+70+10+5) * 100 = 93.33%

Precision

Precision: TP/(TP+FP) * 100

Precision: 80/(80+10)*100 = 88.89%

Recall

Recall: TP/(TP+FN)* 100

Recall: 80/(80+5)*100 = 94.12%

F1-score

F1-score: 2*(Precision*Recall)/(precision+Recall) *100

F1-score = 2 *(88.89% * 94.12%)/(88.89%+94.12%)*100 = 91.43%

Normalized Confusion Matrix

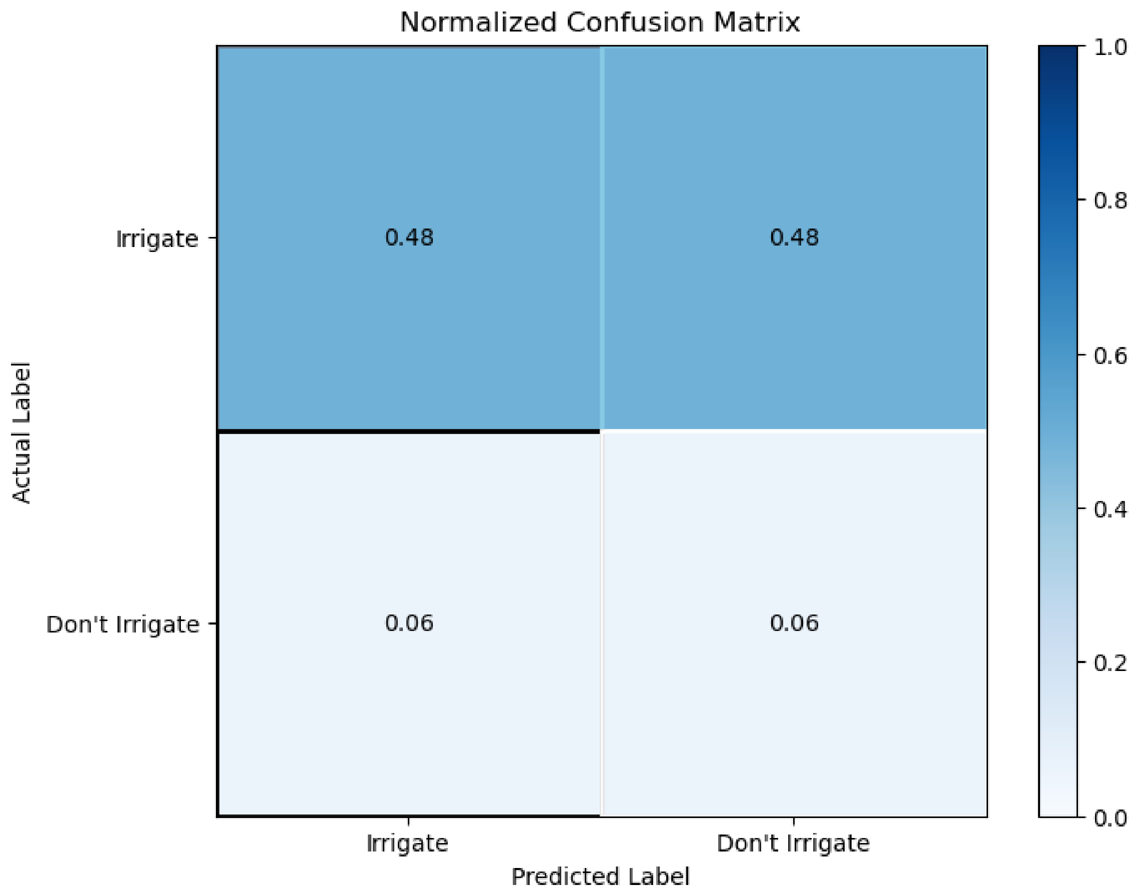

Figure 14: Normalized Confusion Matrix

4.6 Discussion

Advantages

Efficiency: The system can optimize water usage by ensuring irrigation only when necessary, founded on predictive analytics and real-time data.

Scalability: The architecture allows for easy expansion with additional sensors or modules as needed.

Remote Monitoring: The integration with the Blynk platform allows users to monitor and control the system from anywhere.

Challenges

Network Dependency: The system's reliance on network connectivity for communication with the Flask server and Blynk platform could be a limitation in areas with poor internet access.

4.7 Conclusion

The research successfully demonstrated the feasibility and effectiveness of an IoT-based Smart Irrigation System using a Random Forest algorithm for predictive analytics. The system showed high accuracy in predicting irrigation needs, effective integration of hardware and software components, and robust performance in real-time operation. Future work should focus on improving system resilience to network issues and exploring additional features for further optimization. By leveraging IoT technologies and machine learning, the project presents a significant advancement towards sustainable and efficient agricultural practices, providing a foundation for future developments in smart farming solutions.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

The research embarked on developing an IoT-driven Smart Irrigation System aimed at automating irrigation processes through the integration of hardware components and machine learning algorithms. The study unfolded through various stages, including data collection, model training, system integration, and performance evaluation. Key findings from each stage contributed to understanding the efficacy and functionality of the proposed system.

5.2 Major Conclusions Drawn

The research concludes with several key insights derived from the findings:

Data Collection and Generation

Simulated datasets were generated to train the Random Forest algorithm, providing a foundation for predictive analytics in irrigation management.

Model Training

The Random Forest algorithm demonstrated exceptional accuracy, achieving a 100% accuracy rate on the training dataset. This indicated a robust learning capability in capturing patterns and predicting irrigation needs.

System Integration

Integration of the trained model into a Flask web server facilitated real-time prediction and control of irrigation processes, enhancing system automation and efficiency.

Performance Evaluation

The system exhibited reliable performance in controlling the water pump based on real-time sensor data and model predictions, validating its effectiveness in optimizing irrigation.

Effectiveness of Machine Learning:

The high accuracy achieved by the Random Forest algorithm underscores the potential of machine learning in optimizing agricultural practices, particularly in irrigation management.

Automation Benefits:

The Smart Irrigation System demonstrated significant benefits in automating irrigation processes, including improved water efficiency, reduced manual intervention, and enhanced crop yields.

Integration Capabilities:

Seamless integration of hardware components with software algorithms highlighted the importance of IoT technologies in modern agricultural systems, offering scalable and adaptable solutions for farming practices.

Potential for Sustainability:

The implementation of smart farming technologies presents opportunities for sustainable agriculture, minimizing water wastage, reducing environmental impact, and ensuring long-term viability of agricultural ecosystems.

5.3 Recommendations

Based on the research findings, several recommendations are proposed for further consideration:

Real-world Validation:

Future studies should focus on validating the system's performance using real-world data from agricultural settings, ensuring the reliability and applicability of the technology in practical scenarios.

Enhanced Sensor Technologies:

Integration of advanced sensor technologies, such as remote sensing and aerial imaging, could further enhance the accuracy and granularity of data collection, providing deeper insights into crop health and environmental conditions.

Community Adoption:

Efforts should be made to promote the adoption of smart irrigation systems among farming communities, including awareness campaigns, training programs, and incentives for technology adoption.

Continuous Improvement:

Continuous research and development efforts are essential to refine and optimize smart irrigation systems, incorporating feedback from users and stakeholders to address evolving agricultural needs and challenges.

5.4 Conclusion

In conclusion, the research has demonstrated the feasibility and potential of IoT-driven Smart Irrigation Systems in revolutionizing agricultural practices. By leveraging machine learning, automation, and data-driven decision-making, these systems offer tangible benefits in water management, crop productivity, and environmental sustainability. With further advancements and widespread adoption, smart irrigation technologies hold promises in addressing global challenges related to food security, resource conservation, and climate resilience in agriculture.

APPENDICES

Questionnaire

Research: IoT Based Smart Irrigation System using machine learning

My name is Melisa Tariro Kapimbi and I am currently studying for a bachelor's of science degree in Computer Science at Bindura University of Science Education. I am conducting research about the use of Smart Irrigation System in the agricultural industry. The questionnaire consists of 7 questions and will take no longer than 10 minutes to complete. All responses will be kept anonymous, and no one will be identifiable in the research.

Once complete please e-mail back to kapimbitariro@gmail.com.

Please tick the box provided to show your consent to be part of the research. ☐

1. Demographic Information:

   a. Age:

below 18☐ 18-25☐ 25-35☐ 35-45☐ 45-55☐ 55-65☐ >65☐

   b. Gender:

   c. Years of experience in farming:

2. Familiarity with Smart Irrigation Systems:

   a. Have you heard of smart irrigation systems?

      i. Yes

      ii. No

   b. If yes, please describe your understanding or knowledge of smart irrigation systems.

   c. Have you ever used a smart irrigation system on your farm?

      i. Yes

      ii. No

3. Water Management Practices:

a. How do you currently manage irrigation for your crops?

b. What factors do you consider when determining the irrigation schedule?

c. How do you monitor soil moisture levels in your fields?

i. Visual observation

ii. Manual soil moisture probes

iii. Automated soil moisture sensors

iv. Other (please specify)

4. Challenges in Irrigation:

a. What are the main challenges you face when it comes to irrigation?)

b. Have you experienced issues related to overwatering or underwatering of your crops?

i. Yes

ii. No

iii. Not sure

c. If yes, please describe the consequences of inefficient irrigation practices on your farm.

5. Perception of Smart Irrigation Systems:

a. Do you believe that smart irrigation systems can help address water management challenges in agriculture?

i. Yes

ii. No

b. What do you think are the potential benefits of adopting smart irrigation systems?

c. What concerns or reservations do you have about implementing smart irrigation systems?

6. Adoption of Technology:

a. Are you open to adopting new technologies, such as smart irrigation systems?

i. Yes

ii. No

b. What factors would influence your decision to adopt a smart irrigation system?

c. What support or resources would you need to successfully implement a smart irrigation system?

7. Future Perspectives:

a. How do you envision the impact of smart irrigation systems on your farm's productivity and sustainability?

b. What do you believe are the barriers to widespread adoption of smart irrigation systems in your region?

c. What improvements or features would you like to see in smart irrigation systems?

Date: …………..

Location: ………

REFERENCES

Mokyr, J., Vickers, C. and Ziebarth, N.L., 2015. The history of technological anxiety and the future of economic growth: Is this time different?. Journal of economic perspectives, 29(3), pp.31-50.

Mubvami, T., Mutenje, M. J., & Dube, T. 2020. Review of Machine Learning Techniques for Irrigation Water Demand Forecasting in Agriculture. Agronomy, 10(5), 751.

M. A. Gumbe, G. Ndengu, and T. M. Mushore, "Climate variability and change trends in Zimbabwe," International Journal of Climate Change Strategies and Management, vol. 8, no. 3, pp. 338-356, 2016; R. Ncube and J. Ncube, "The contribution of agriculture to Zimbabwe's economy: Facts and myths," Cogent Economics & Finance, vol. 6, no. 1, 2018.)

Patil, R. J., Mulage, I., & Patil, N. (2023). Smart agriculture using iot and machine learning. *Journal of Scientific Research and Technology*, 47-59.

Heermann, D. F., & Heinemann, A. B. (2019). Principles of Irrigation. In Handbook of Water Resources Management: Discourses, Concepts and Examples (pp. 267-297). Springer.

Islam, M. R., Islam, M. N., Kawsar, M. H., Islam, M. A., & Kundu, D. K. (2019). Importance of Irrigation in Agriculture. Agriculture and Food Sciences Research, 6(3), 235-238

Food and Agriculture Organization. (2021). Zimbabwe Crop and Food Security Assessment Mission.

Archana and Priya, "Design and Implementation of Automatic Plant Watering System" presented at International Journal of Advanced Engineering and Global technology , vol-04, Issue-01 , Jan-2016.

C.H.Chavan and V.Karnade ," Wireless Monitoring of Soil moisture, Temperature and Humidity using Zigbee in Agriculture" presented at International Journal of Engineering Trends and Technology (IJETT) ,vol-11, May-2014.

Sonali.D.Gainwar and Dinesh.V.Rojatkar ,"Soil Parameters Monitoring with Automatic Irrigation System" presented at International Journal of Science, Engineering and Technology Research(IJSETR),vol- 04,Issue 11,Nov 2015.

Abdul A. I., Hasan M. H., Ismail M. J., Mehat M., and Haroon N. S., "Remote Monitoring in Agricultural Greenhouse Using Wireless Sensor and Short Message Service (SMS)", In Proceedings of International Journal of Engineering & Technology IJET-IJENS, Volume 9, No 9, PP 1-9, October 2009.

https://www.fao.org/agriculture/crops/thematic-sitemap/theme/irrigation/en.

Okoli, C. N. (2015). How to write a literature review. Interfaces, 45(2), 27-40. doi:10.1287/inte.2014.0990.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 29(7), 1645-1660. doi:10.1016/j.future.2013.01.010.

Food and Agriculture Organization of the United Nations (FAO). (2019). Zimbabwe drought response. Retrieved from https://www.fao.org/emergencies/drought-in-zimbabwe/overview/en/.

Lema, I. G., Nyamangara, J., Jakobeit, H., Mushore, D., & Dube, A. (2019). Assessing the state of water shortages for irrigation development in the Sanyati Catchment, Zimbabwe. International Journal of Water Resources Development, 35(2), 301-315. https://doi.org/10.1080/07900627.2019.1549447

Marinda, S. (2016). Smallholder irrigation technologies in Zimbabwe. Renewable and Sustainable Energy Reviews, 65, 387-398. https://doi.org/10.1016/j.rser.2016.07.032.

Ilemobade, T. O. (2017). Solar-powered drip irrigation technology for smallholder farmers in the tropics: A case of Zimbabwe. Renewable and Sustainable Energy Reviews, 79, 773-783. https://doi. org/10.1016/j.rser.2016.10.127.

IBM Watson. (n.d.). What is machine learning? Retrieved from https://www.ibm.com/topics/machine-learning

P. Suguna, & M. Prakash. (2018). Automated irrigation system for agricultural field using GSM technology. International Journal of Advanced Research

Gogoi, M. K., Kar, B., Saha, M., & Patgiri, A. (2018). Automatic irrigation system using soil moisture sensor with remote monitoring. Journal of Emerging Trends in Engineering and Applied Sciences, 9(11), 546-553. Retrieved from https://www.scribd.com/document/370305958/Automatic-irrigation-system-using-soil-moisture-sensor-with-remote-monitoring.

Saunders, M., Lewis, P., & Thornhill, A. (2016). Research methods for business students (7th ed.).Pearson Education

Turnitin Report

Tariro_Dissertation.docx

ORIGINALITY REPORT

10%	7%	3%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Midlands State University Student Paper	1%
2	dokumen.pub Internet Source	1%
3	Submitted to University of Greenwich Student Paper	<1%
4	ieomsociety.org Internet Source	<1%
5	open-innovation-projects.org Internet Source	<1%
6	Jinyuan Xu, Baoxing Gu, Guangzhao Tian. "Review of agricultural IoT technology", Artificial Intelligence in Agriculture, 2022 Publication	<1%