

## Paddy Rice Smart Farming

Hakkin Chethiya Kaushila De Silva<sup>1</sup>, Srisinthujan Shanmuganathan<sup>2</sup>, Mathusan Anantharajah<sup>3</sup>, Sivanujan Sivanathan<sup>4</sup>,  
Thamali Dassanayake<sup>5</sup> and Prof. Sanath Jayawardena<sup>6</sup>

<sup>1</sup>Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, SRI LANKA

<sup>2</sup>Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, SRI LANKA

<sup>3</sup>Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, SRI LANKA

<sup>4</sup>Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, SRI LANKA

<sup>5</sup>Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, SRI LANKA

<sup>6</sup>Department of Computer Systems Engineering, Sri Lanka Institute of Information Technology, Malabe, SRI LANKA

<sup>2</sup>Corresponding Author: it19217550@my.sliit.lk

### ABSTRACT

It is anticipated that machine learning (ML) and the internet of things (IoT) would significantly impact smart farming and engage the entire supply chain, in particular for the production of rice. Rice smart farming offers new capabilities to foresee changes and find possibilities thanks to the growing amount and variety of data gathered and obtained by emerging technologies in the Internet of Things (IoT). The accuracy of the models created through the use of ML algorithms is significantly impacted by the quality of the data obtained from sensor readings. These three components, machine learning (ML), the internet of things (IoT), and agriculture have been used extensively to improve all aspects of rice production processes in agriculture. As a result, traditional rice farming practices have been transformed into a new era known as rice smart farming or rice precision agriculture. We do a study of the most recent research that has been done on the application of intelligent data processing technology in agriculture, namely in the production of rice, in this paper. We analyze the applications of machine learning in a variety of scenarios, including smart irrigation for paddy rice, predicting paddy rice yield estimation, predicting droughts and floods, monitoring paddy rice disease, and paddy rice sample classification. In each of these scenarios, we describe the data that was captured and elaborate on the role that machine learning algorithms play in paddy rice smart agriculture. This paper also presents a framework that maps the activities defined in rice smart farming, data used in data modeling, and machine learning algorithms used for each activity defined in the production and post-production phases of paddy rice.

**Keywords--** AI, Deep Learning, IoT, Machine Learning

### I. INTRODUCTION

Floods do the most damage and are notoriously hard to predict. Improved flood prediction models helped reduce risks, propose regulations, and minimize human and property loss from floods [1]. Machine learning (ML) approaches have led to increasingly accurate and cost-efficient prediction systems during the past two decades. ML's many advantages and prospects have boosted its appeal among hydrologists. Researchers want

to produce more accurate and efficient prediction models by hybridizing machine learning technologies. This research shows the present status of machine learning models for flood prediction and identifies the best models. This publication provides a comprehensive overview of machine learning algorithms [1]. To accomplish so, we analyzed ML model research for robustness, accuracy, efficacy, and speed. The performance comparison of ML models provides a comprehensive grasp of the various approaches. This study offers the most promising long-term and short-term flood prediction approaches. Also studied are the major breakthroughs in improving flood prediction models. Hybridization, data deconstruction, algorithm ensemble, and model optimization are said to improve ML techniques. The most effective is hybridization.

Irrigation may be required during dry periods in order to maintain the quality of the landscape [3]. It is possible to increase the incidence of disease in a landscape, wastewater, and worsen the landscape's general state if the landscape is either over- or under-irrigated. The effectiveness of an irrigation system is determined by a number of elements, including the design, the installation, and the particular characteristics of the location. The amount of water that is used to maintain a property's landscaping might make up a sizeable fraction of the total amount of water used on the site. In Paddy, household water consumption for outdoor uses represents for around 30% to 50% of total household water consumption [3]. As a direct result of the use of inefficient watering practices, a sizeable quantity of water is lost as a result of evaporation, wind, and runoff. This loss, if reduced or eliminated entirely, can result in lower utility bills and the creation of a healthy landscape that uses less water. Utilizing efficient irrigation methods can help achieve water savings in outdoor spaces. In comparison to standard automatic system timers, which irrigate on a user determined set schedule, smart irrigation controllers and sensors have been created to reduce the amount of water used for outdoor irrigation by irrigating plants based on the amount of water they require. This technology is available in the form of a complete controller as well as a sensor that can be attached to a conventional irrigation

timer in order to transform it into a smart controller. The level of irrigation that a landscape requires can be determined with the use of weather data or data on the moisture content of the soil using smart irrigation technology.

Rice is one of the foods that contributes the most to the global economy in this day and age. According to a report by the FAO, there are a variety of reasons why it would be desirable to manage rice processing and the byproducts that result into applications that are more environmentally friendly. In order to generate consumable products at the end of the processing of rice, there are multiple milling processes. The milling process is the most significant phase in the manufacture of rice because it is responsible for determining the nutritional value, the ability of the rice to be cooked, and its flavor. As rough rice goes through the milling process, byproducts such as bran are produced. These byproducts, which have been demonstrated to have positive effects on the nutrition of both humans and animals, are called bran. Rice bran is one of the many rice byproducts that can be used in agriculture, but it is perhaps the one that has garnered the most interest due to the useful features it possesses. In order to do this, we examine by-product prediction with a model that makes use of an Artificial Neural Network.

Machine learning, deep learning, and a great many other methods are among those that are utilized to simplify agricultural practices and assist farmers in achieving higher levels of productivity [8]. An application that categorizes leaf diseases, as well as recommends appropriate crops and fertilizers, has been developed by our team. Leaf disease is one of the most significant ailments that can affect paddy. Due to the lack of access to specialists, diagnosing paddy leaf diseases can be an extremely time-consuming and tedious process for farmers in distant places. Even if there are specialists accessible in certain fields, disease detection is still done with the naked eye, which might lead to incorrect recognition. These issues can be reduced to a manageable level by using an automated system. In this article, an automated approach is provided for the identification of three prevalent paddy leaf diseases: brown spot, leaf blast, and bacterial blight. Based on the severity of the illnesses, recommendations are made for the use of pesticides and/or fertilizers.

Therefore, this study is about the implementation of a Supporting application to Enhance Paddy rice smart farming and can be implemented using multiple machine learning algorithms and classification techniques to provide the highest accuracy and best solutions for user to improve their performance skills.

## II. BACKGROUND

In past years, several systems have been proposed for Paddy rice smart farming. Different machines for Paddy rice smart farming have been used.

Below are some articles we have reviewed on the Paddy rice smart farming application.

Using machine learning and neural networks, a hybrid model was created to predict floods in Sri Lanka's North Central Province. The hybrid model combines two sub predictive models. The first model anticipates weather using time series modeling. The second model, a binary classification machine learning system, forecasts the probability of future flood occurrences using meteorological forecasts and previous data. This model uses previous data. The hybrid model's predicted probability have a correlation of 91.7% or higher with actual flood occurrences. Using historical meteorological data and flood data, this model can predict flooding in any portion of Sri Lanka. The produced prediction model has been provided as an API in Microsoft Azure's cloud, showcasing the practical application of machine learning techniques in the building of intelligent apps. This method can forecast users' continued use of the home workout application after installation. On Keep, they took into account 18 variables that were predicted to have an impact on the determination of continuous participation. Keeping certification was the main factor influencing the study's findings.

It was proposed that a study been conducted in which humidity and soil moisture sensors would be implanted into the root zone of the plant. On the basis of the data that was measured, the microcontroller is utilized to regulate the amount of water that is available in the sector. Because it does not provide information on the state of the sector, this method is useless for farmers and is therefore ineffective. [4] A document was suggested that incorporates the utilization of photovoltaic cells for the purpose of obtaining power. Electricity is not necessary for the operation of this system. In order to turn the motor pump on and off, a soil moisture sensor is utilized, and a PIC microcontroller that is supported by the measured data is utilized. The field of meteorology is excluded from this system.

In this research [5], they examine meteorological and soil data-based yield estimation modeling for paddy crop at different elevations. India has coarse (district) and fine (taluk) spatial resolution levels. We analyze the accuracy of yield estimation models across features and ML approaches. Disaggregating district yield data involves utilizing machine learning models trained on district data to predict taluk yields. Disaggregating district-level data to forecast taluk yield yields a 6% average error and a 25% maximum inaccuracy.

CNN is used to construct the proposed system [6]. A pretrained ResNet-50, ResNet-101, VGG-16, VGG-19, EfficientNet, Inspection-V2, and GoogleNet library are used in the construction of CNN. To improve the accuracy of the identification process while maintaining its high level of productivity, a ReLU classifier is utilized. This model may assist the farmer in recognizing the condition of the paddy leaf as a key diagnostic tool, and it can also assist the agriculturist in

confirming his prediction by inspecting the paddy leaf. The standard operating procedures in the laboratory are both expensive and time-consuming. The rice leaf disease detection method that is proposed in this research is going to identify and diagnose five different categories of paddy leaf diseases. [8] The detection, classification, and diagnosis of paddy leaf diseases such as rice blast, brown spot, and bacterial blight can be predicted with the help of a k-means clustering model. For the purpose of classifying the paddy leaf illness, procedures such as picture acquisition, image preprocessing, image segmentation, disease detection, transformation to a binary image, and calculation of the affected percentage area of leaf were utilized.

By considering these factors we recommend a Mobile application as a solution and hope it will be of help as a Paddy rice smart farming application for farmers.

### III. METHODOLOGY

Fig.1 shows a block diagram of the Integrated Smart paddy system: A efficiency application for Enhancing the paddy field AI system. The application's primary goal is to enable support farmers to get a high-yield system. The mobile application will aid in Agriculture.

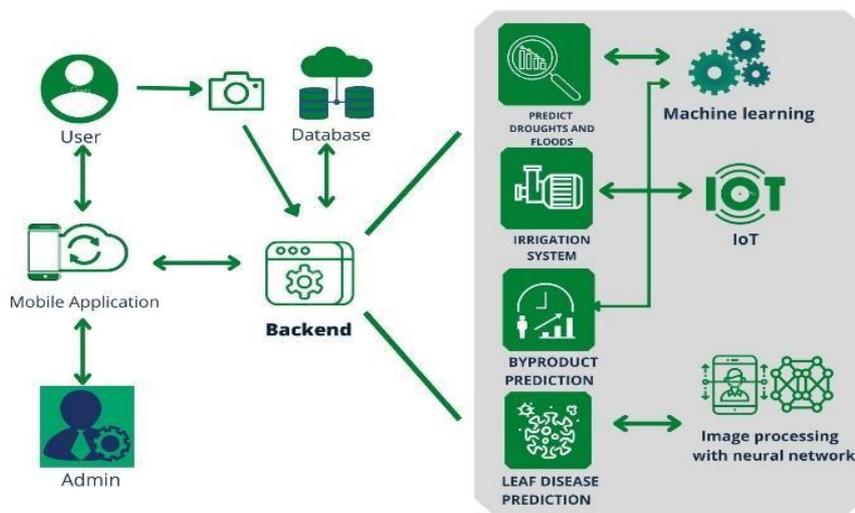


Figure 1: System Architecture

The Proposed system will be carried out under four main components.

- 1) Predict possible droughts and floods Using ML
- 2) Smart Irrigation System to regulate the water flowing IoT.
- 3) Predict the amount of Byproduct obtained in a paddy yield cultivation
- 4) Paddy leaf disease prediction Using Image processing.
- 5) Give the solution to build a Smart Paddy farming system

- Floods that may occur in the future can be accurately predicted here.
- Determine the likelihood of future droughts and accurately forecast them.
- Through the use of a mobile application, farmers should be informed about these incidents

Secure and user-friendly application system based on machine learning application using 4 types of models. The program will guide farmers to overcome the given issues.

Different types of CNN algorithms and models will be used to examine the data. After the ML (Machine Learning) model has been trained, the binary classifier results will be categorized. Python was chosen as the programming language, along with the libraries, Deep Learning, ML framework, K-means clustering Model to Implementation, and Image processing for finding the disease.

#### A. Predict possible droughts and floods Using ML

In order to prevent rice fields from being destroyed by floods and droughts, this paper develops a system that uses techniques from machine learning and communicates this information to the farmers.

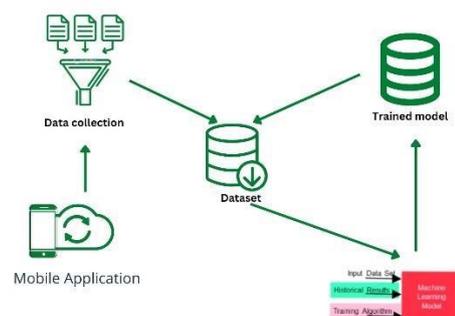


Figure 2: Predict possible droughts and floods

Above Fig. 2 is the individual system diagram where this component is represented as a small box, and the middlebox in the research component is typical for everyone; hence it involves machine learning. It is essential for the long-term planning and management of water resources at both the global and regional levels to accurately forecast meteorological and hydrological droughts using standardized measures of rainfall and run off (SPI/SRI). In this study, several machine learning (ML) techniques, including four methods (i.e., ARIMA, SARIMA, FBP, and LSTM), were used to construct hydrological drought forecasting models in the Akkarapathu, Hingurakgoda, and Nikaveratiya basins in Sri Lanka. These models were used to predict the likelihood of drought conditions.

In this study, we (a) identify concerns that are similar to the prediction of drought and floods and their simultaneous assessment, and (b) explain various strategies that may be implemented to overcome these challenges. The difficulties associated with forecasting floods and droughts can be broken down into four distinct but interconnected categories: data, understanding of processes, modeling and prediction, and human interactions with water. The availability of data and the definition of events are both challenges linked with data. Extremes' multivariate and spatial features, non-stationarities, and future alterations are process-related concerns. Frequency analysis, stochastic modeling, hydrological modeling, earth system modeling, and hydraulic modeling present distinct modeling challenges. Human–water interactions provide obstacles in establishing impacts, adequately describing interactions, and conveying scientific findings.

**a) ARIMA**

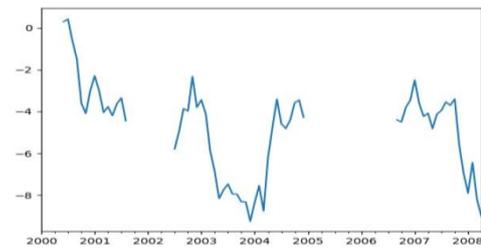
ARIMA models add to time series forecasting. Exponential smoothing and ARIMA models are used primarily in time series forecasting. These approaches clash with one another. ARIMA models strive to capture data autocorrelations, whereas exponential smoothing techniques are based on data trend and seasonality. A period series is fixed when its properties don't vary over time.

ARIMA models could provide some potential benefits.

- In order to generalize the forecast, one only needs to use the previous data of a time series.
- Performs well in terms of expectations for the near future.
- Models time series that are not stationary.
- ARIMA models may have several potential drawbacks.
- It is difficult to anticipate where the turning points will be.

**b) SARIMA**

SARIMA uses historical values and seasonality patterns similarly. SARIMA is more powerful than ARIMA in forecasting complex data cycles because it includes seasonality as a parameter.



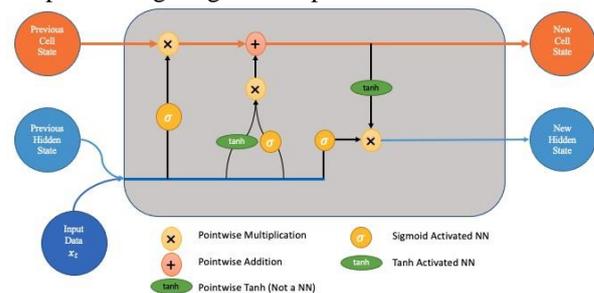
**Figure 3: SARIMA**

**c) FBP Model**

Starting with the iterative Landweber algorithm, a frequency domain window function is derived for each iteration. The resultant window function emulating the Landweber algorithm after  $k$  iterations have a  $k$  index. The FBP window function adjusts the ramp filter. FBP creates CT scans quickly and easily. According to the name, it consists of two steps: filtering the data (in the detector's row direction) and back projection or painting the data back in the image along the direction it was measured. Both steps are sequential.

**d) LSTM**

For the sake of this study, we decided to use LSTM because it has consistently shown itself to be one of the most effective algorithms for producing accurate forecasts of time series. It is well known that the LSTM model is good at remembering long-term dependencies. Even if we could suppose that the weather conditions from the previous day are all that is needed to accurately forecast today's rainfall or any other weather condition, this line of thinking is illogical. Because an event occurs at a specific time accompanied by a sequence of changes, this is so. Because of the presence of a memory cell and a reset cell inside its structure, LSTM is capable of comprehending long-term dependencies.



**Figure 4: LSTM**

**B. Smart Irrigation System to regulate the water flowing IoT**

This system was assembled using a Raspberry Pi, a soil moisture sensor, a DHT 11 sensor, an ultrasonic sensor, a buzzer, and an MQ-04 methane gas sensor. The Raspberry pi serves as the system's principal controller and is hence considered its "heart." This system was created using Raspberry Pi, which has a number of modernized features. It includes greater input/output (IO) and connectivity when compared to previous generations of the Pi. The moisture content of the soil is intrinsically

tied to the growth of raspberry pi. The information gathered by the sensor is transmitted to the Raspberry Pi as soon as it is detected. The Raspberry Pi's behavior will be determined by the information it receives from the soil moisture sensor. If the soil is dry, meaning the sensor is not detecting any moisture, it will immediately engage the pump connected to the irrigation system. When the water level reaches the point, regardless of whether the motor has been started or not, it will automatically shut off. The DHT 11 Sensor is required to detect the presence of water on a farm. The DHT-11 Sensor collects environmental data, which is subsequently transferred to the Raspberry Pi.

Consequently, if there is water, we simultaneously set off the alarm on the buzzer. Using an ultrasonic sensor, we are determining how much water is there in the well. Now that we have the farmers' aid, we can accurately determine the water level by measuring the well's depth and its circumference. Now, with the aid of an ultrasonic sensor, we will measure the current water level in the well, compute the water level beginning from the bottom of the well, and calculate the amount of water in the well at this time.

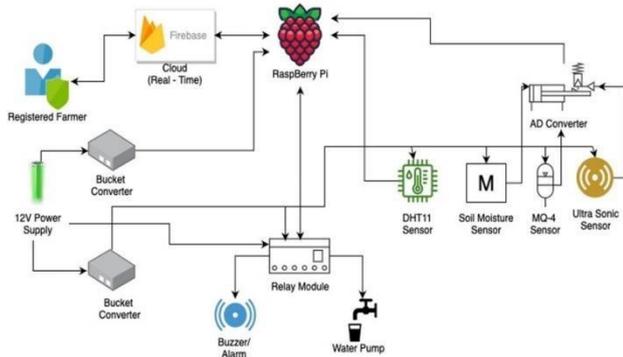


Figure 5: Smart Irrigation System

The availability of water is a critical factor in paddy crop production. During the whole growing season, the field should have a constant supply of standing water. Rice thrives on flooded soil due to its favorable features, which are as follows:

Rice crops benefit from submergence because it helps to restrict the growth of weeds and makes specific nutrients more accessible.

- The rice crop's daily water uses ranges from 6 to 10 millimeters (mm).
- A total amount of water ranging from 1200 to 1400 millimeters is required for a rice crop.
- To produce 1kg of rice, 2000 - 3000 liters of water are required.
- Highly saline and brackish water not suitable for rice crop irrigation.

C. Predict the amount of By-product obtained in a paddy yield cultivation

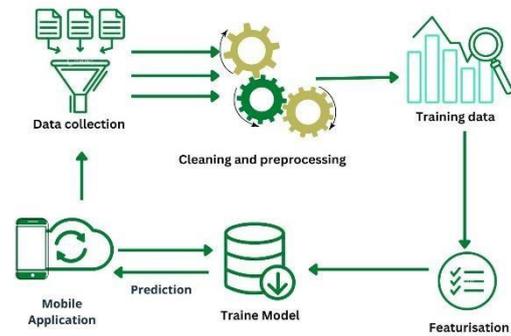


Figure 6: Predict the amount of By-product

To forecast the amount of by-product as well as to provide suggestions for maximizing harvest rate while minimizing by-products. Prepare the best machine learning model possible to make a by-product prediction for paddy crops. Find some ways to cut down on waste while increasing the harvest rate. The figure shows the current process of the system. Once a user logs in to the system, they can view the variety of rice. Then they may select the rice that they need. The system will provide information about the selected product. Once the system is done, a trained model will search for the product prediction to give rice. Finally, the most accurate prediction will be shown to users with a recommendation.

a. Random Forest

The Random Forest meta-algorithm is one that makes use of a number of different Decision trees. It does this by combining the prediction values of all of the decision trees and using a method called majority vote, which returns the class with the majority of the votes. Occasionally, the Decision tree will grow quite deep, at which point it will encounter the problem of overfitting and learning unusual patterns. However, the random forest provides a solution to this overfitting issue.

The randomness of the Random Forest classifier is what makes it so famous. It delivers unpredictability in two different ways, the first being randomness connected to the data, and the second being randomness related to the characteristics. Both Bagging and Bootstrapping are concepts that are utilized by the Random Forest classifier.

- Characteristics of the Random Forest
- The random forest has two characteristics, which can be broken down as follows:

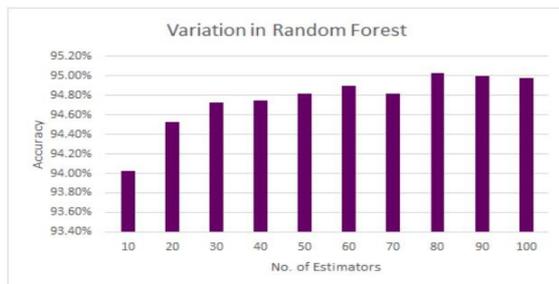
**Robust:** Random forest is a robust machine learning technique because it averages the output of several decision trees, each of which was trained using a unique set of data.

The term for this practice is "bootstrapping." In this scenario, every tree looks at a different, randomized sample of the training data. Therefore, it is significantly more resistant to noise.

**Accuracy:** the random forest algorithm makes use of the bagging concept. In this, the final output is determined by taking an average of all of the classifiers.

If you send an enormous amount of data to a single classifier, you won't get an accurate result. However, if those data can be divided into a number of classifiers, then the average of those classifiers' results will give you a consistent solute on the whole.

According to our findings, random forest demonstrates a high level of accuracy and results for all of the dataset splitting ratios. The performance of a random forest is contingent upon a large number of parameters, but the number of estimators that are employed is the primary factor. The accuracy of random forest also grows proportionally with the number of decision trees that are used. The result of this is that the variation in random forest caused by varying the number of estimators is shown in figure. as well as a graphical representation of it that may be found below:



**D. Paddy leaf disease prediction Using Image processing**

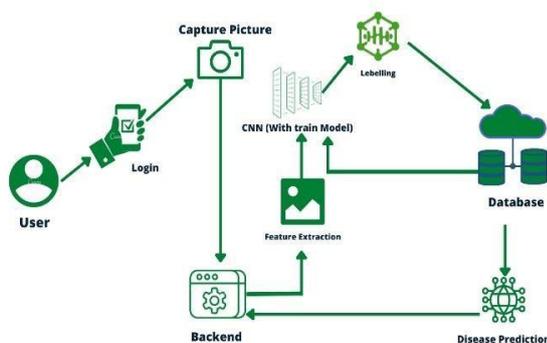


Figure 7: Paddy leaf disease prediction

If a farmer wants to evaluate a leaf that has been affected by a disease and needs to be certain that the plant is suffering from a specific disease or not, then all he needs to do is use the mobile application and take a picture of the disease to determine whether or not the plant is afflicted with the disease. a leaf of the paddy that is afflicted.

These are the selected disease Brownsport, rice blast, Bacterialblight, tungro.

**a. CNN**

The performance of several distinct deep convolutional neural network (DCNN) models that varied in the number of convolutional (Conv) and pooling layers as well as their sizes has been analyzed. Depending on the specific DCNN model, the number of Conv layers can range anywhere from three to eight. When trained to its full potential, the 14-layered deep convolutional neural network (14-DCNN) provides superior performance to that of other models that have been developed. The suggested 14-DCNN model was built using a total of ten layers, including five convolutional and five max-pooling layers. The first two-dimensional convolutional layer of the 14-DCNN receives the input images from the previous layer. Utilizing Equation (1)[9], one can determine the size of the output from the Conv layer as follows:

$$Dimension(Conv(n,k)) = \left( \left[ \frac{n_w - f_w}{s} + 1 \right], \left[ \frac{n_h - f_h}{s} + 1 \right], f_c \right)$$

Both the width (nw) and height (nh) of the input to the first convolutional layer are also set to 128. In addition, the width, height, and channels of the kernel filter of the convolutional layer are represented by the fw, fh, and fc variables, respectively. This Convolution layer has a value of one for its stride (S) parameter. In order to decrease the dimension of the output of the first Conv layer from 126, 126, 4 values to 63, 63, 4 values, the first max-pooling layer was developed. The dimension of the output from the maximum pooling layer was determined by utilizing Equation (2)[9]:

$$Dimension(Pooling(n,k)) = \left( \left[ \frac{n_w - f_w}{s} + 1 \right], \left[ \frac{n_h - f_h}{s} + 1 \right], n_c \right)$$

The suggested model employing CNN was educated on data taken from the paddy leaf disease dataset. It was able to achieve an accuracy of 94.87% while recognizing infected plants with the assistance of image processing performed by OpenCV. In conclusion, the research study presents a comprehensive examination of the proposed scheme as well as the findings acquired by the model.

**IV. RESULTS**

A series of classification reports, confusion metrics, loss, accuracy graphs, and the roc curve has been developed in order to evaluate the validation accuracy of the rice disease prediction. This was done so that the curve could be used. Accuracy, precision, recall, and f1 score are the major metrics that have been taken into consideration for the purpose of measuring performance in this study. A confusion matrix was developed so that the level of accuracy achieved by the model's validation could be demonstrated. It has been determined that a 95% accuracy has been achieved. The Decision Tree Algorithm can be used for future predictions at a low computational cost. The loss curve depicts the network's

learning process. Bacterial blight, rice blast, and brown spot all had 98% classification accuracy. Climate change has a large impact on paddy byproducts. A study concludes that merging IoT and automation technology can boost the irrigation business.

This article shows how to use a random forest classifier to control irrigation. The suggested system can turn the on and off the sprinkler in response to changes in soil moisture, simplifying operation.

## V. CONCLUSION

The primary focus of this study is on Agriculture which is Sri Lanka's main industry. Climate change affects agriculture harvesting in Sri Lanka. Failures in weather forecasting, catastrophe management, and disease control affect agricultural output. As a solution, a decision support system was created. This technique targets Sri Lanka's Western Province, a wet zone low country, but it can be used in other provinces and biomes. Rice disease can also affect other crops.

## REFERENCES

- [1] A. Mosavi, P. Ozturk & K. Chau. (2018). Flood prediction using machine learning models: Literature review. *Water*, 10(11), 1536. DOI: 10.3390/w10111536.
- [2] K. Sundararajan *et al.* (2021). a contemporary review on drought modeling using machine learning approaches. *Computer Modeling in Engineering & Sciences*, 128(2), 447– 487. DOI: 10.32604/cmescs.2021.015528.
- [3] M. D. Dukes. (2009). Smart irrigation controllers: what makes an irrigation controller smart?. *EDIS*, 2009(2). DOI: 10.32473/edis-ae442-2009.
- [4] V. K. S, C. D. Singh, K. V. R. Rao, M. Kumar & Y. A. Rajwade. (2022). Development of a smart IoT-based drip irrigation system for precision farming. *Irrigation and Drainage*. DOI: 10.1002/ird.2757.
- [5] A. R. Bodie, A. C. Micciche, G. G. Atungulu, M. J. Rothrock & S. C. Ricke. (2019). Current trends of rice milling byproducts for agricultural applications and alternative food production systems. *Frontiers in Sustainable Food Systems*, 3. DOI: 10.3389/fsufs.2019.00047.
- [6] Md. E. Farooqui & Dr. J. Ahmad. (2020). Disease prediction system using support vector machine and multilinear regression. *International Journal of Innovative Research in Computer Science & Technology*, 8(4). DOI: 10.21276/ijirest.2020.8.4.15.
- [7] O. K. Pal. (2021). Identification of paddy leaf diseases using a supervised neural network. *16th International Conference on Emerging Technologies (ICET)*. DOI: 10.1109/icet54505.2021.9689788.
- [8] S. Nalini *et al.* (2021). Paddy leaf disease detection using an optimized deep neural network. *Computers, Materials & Continua*, 68(1), 1117–1128. DOI: 10.32604/cmc.2021.012431.

- [9] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk & D. Stefanovic. (2016). Deep neural networks based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience*, 2016, 1–11. DOI: 10.1155/2016/3289801.