Crop Prediction and Plant Disease Detection Using Machine Learning

Ms. Anamika Wasnik[1], Mrs. Sinu Nambiar[2] ,Rajvi Deshmukh[3], Anuja Marwadkar[4], Nidhi Meena[5]

[1,2,3,4,5] *Dept. of Computer Science Engineering*
[1,2,3,4,5] *Dr. D.Y. Patil Institute of Technology, Pune, Maharashtra, India*

Abstract - *In general, India's economy relies heavily on agriculture, which also contributes a significant amount of the country's gross domestic product to the nation's efforts to secure food security. Yet, due to manmade climatic changes, food production and forecasting are currently declining, which will have a negative impact on farmers' economies by resulting in a low yield and make farmers less adept at predicting future crops. By utilizing machine learning, one of the most cutting-edge technologies in crop prediction, this research aids the novice farmer in a way that directs them towards sowing the reason- able crops. A supervised learning algorithm offers a strategy for doing so. Here, the appropriate parameters for temperature, humidity, and moisture content are collected along with the seed data of the crops, assisting in the successful growth of the latter. An important agricultural export commodity from Thailand, the Barracuda mango (Nam-Dok Mai), was the subject of this study on the development of an expert system for diagnosing plant diseases. Yet, due to Thailand's tropical climate, several plant diseases have been found there that have an impact on the survival of mango trees. Several different sorts of agricultural production are diminished when an agriculturalist is not aware of how plant illnesses should be categorized. Furthermore, there is no framework in place to suggest the best way to prevent or cure the illness that develops on their farm. As a result, their treatments for sick plants suffer substantially.*

Key Words*: *Precision farming, machine learning, forecasting, categorization, detection, and plant diseases and pests.

1. INTRODUCTION

For a very long time, it has been believed that the primary source of supplies for supplying people's basic needs has been agriculture. It is recognized as both a necessary kind of employment and a significant industrial sector in India. In order to maintain a healthy diversity, farmers should engage in traditional naked eye observation and produce healthy crops without using pesticides on their cultivation field or on the animals that eat such products. However, in today's world, weather patterns are shifting swiftly in opposition to the availability of natural resources, which decreases food supply and increases security. Farms as an industry will have an impact on farmers' way of life because they live in rural locations where crop production revenue is impacted. The agriculture sector's contribution to the GDP was roughly 17.211.1 in 2005 and only about 52020 in 2018. A variety of plant leaf disease images are used in this study's deep convolutional network model to enable rapid and precise automated identification. Symptoms of plant leaf diseases might vary. Plant pathologists with more experience may be able to spot infection earlier than farmers with less training. An autonomous system created to recognize crop illnesses by their look and visual symptoms as a confirmation mechanism in disease diagnosis would be very helpful to farmers. It has required a lot of work to create quick and reliable methods for diagnosing leaf diseases. Plant leaf disease can be identified using neural networks and digital image processing methods. Over the past few years, deep learning has made remarkable strides. Now, useful feature representations can be extracted.

2. LITERATURE SURVEY

Not only does agriculture play a significant role in the expanding economy, but it is also vital to our survival. It is difficult to predict agricultural output since it depends on a variety of factors, including water, ultraviolet (UV) radiation, pesticides, fertilizer, and the amount of land that is covered in that region. Two distinct Machine Learning (ML) techniques are

suggested in this paper to analyze crop yield. Support Vector Regression (SVR) and Linear Regression (LR) are two techniques that are well suited for verifying the variable parameters in the prediction of continuous variables using the 140 data points that were collected. The elements listed above have a significant impact on crop output. Mean Square Error (MSE) and Coefficient of Determination (R2) were used to calculate the error rate; MSE provided a value of about 0.005 and R2 provided a value of about 0.85. The same dataset has been utilized to quickly compare the performances of the algorithms.

Due to a variety of factors, academics have recently developed an interest in land classification and mapping. The growing need for agricultural land and soil health analysis, which are both crucial for healthy crop production, are the causes of the scientific community's increased focus. One method for analyzing the health of soil and land is image classification. It is a complicated procedure with multiple impacts. Examining previous research, the problems it has addressed, and its future directions has been recommended in this essay. The main objective is the analytical analysis of many novel and efficient categorization mechanisms and methods. It has been attempted to study the components that various approaches have addressed in order to improve the classification's accuracy. The selection of the best classifier and the effective use of the multiple features present in remotely sensed data are the two elements that are most important for improving classification accuracy. For the classification of multi-source data, knowledge-based classification, or non-parametric classifiers such as decision trees or neural networks have recently grown in favor. More study is still needed, though, to reduce uncertainties and increase the precision of the image categorization algorithms.

The economy of our country is significantly impacted by the agricultural sector. The main driver of civilizational growth has been agriculture. India is a largely agricultural country with a crop-based economy. We can contend that agriculture can sustain the economy of our nation as a result. When organizing an agricultural project, each crop must be properly chosen. The choice of crops will be influenced by a variety of factors, including market price, production rate, and government policies. To improve changes in our Indian economy, the agriculture sector needs to undergo numerous modifications. Using machine learning techniques that are simple to use in the farming industry, we can improve agriculture. Accurate and helpful information on many issues is essential, in addition to all the advancements in the farming gear and technologies. The objective of this study is to apply the crop selection strategy in order to assist farmers and agriculturalists in resolving a range of problems. As a result, crop yield rates are increased, which is good for the Indian economy.

3. METHODLOGY

The recommended method notifies the farmer of crop diseases so they can take further action. As soon as an illness starts to spread to a leaf's outer layer, the proposed technique tries to identify it. The use of an automatic technique for plant disease detection has several advantages, including the ability to detect disease symptoms at an incredibly early stage, such as when they first appear on plant leaves, and the reduction of labor needed to monitor vast farms of crops. The proposed method provides a method for routinely monitoring the cultivated area and a mechanism to automatically detect plant leaf disease using remote sensing photographs. The model can be changed so that it works with the images

rather than the dataset. With the deep learning technique, which will cover more minute features, the model can be made to be more effective.

4. ARCHITECTURE DIAGRAM

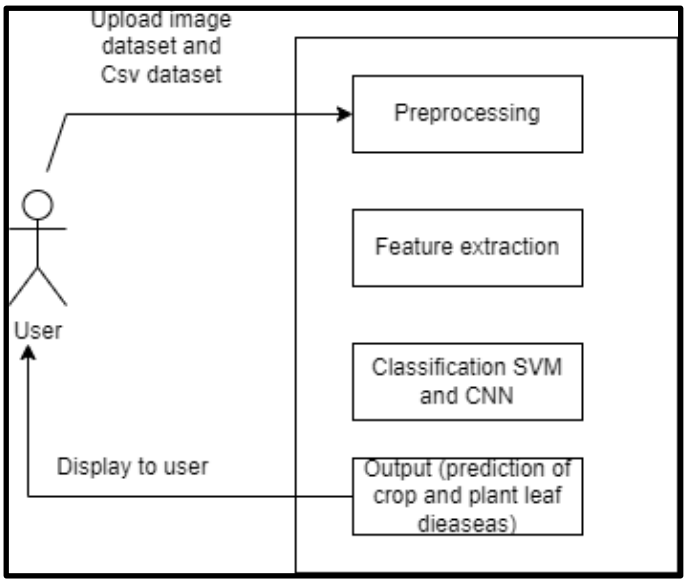

Fig 4.1. Architecture Diagram

5. ALGORITHM

1. CNN Algorithm:

A Deep Learning system called a convolutional neural network may rank several portions of an input image according to their significance and discriminate between them.

The CNN-RNN is unique in three ways that make it a potentially useful method for future studies on agricultural yield prediction. (1) The CNN-RNN model was developed to reflect how environmental variables change over time and how genetically diverse seeds develop. (2) The model demonstrated that there is no loss in forecast accuracy when the yield prediction is applied to untested scenarios. (3) The model may demonstrate how much the weather, the accuracy of weather predictions, soil conditions, and management practises may contribute to the variation in crop yields when used in conjunction with the backpropagation approach. It can be trained to recognise different elements in the photos, including the size, colour, texture, and form of the crops, in order to determine the type of crop, calculate crop production, and keep track of crop health.

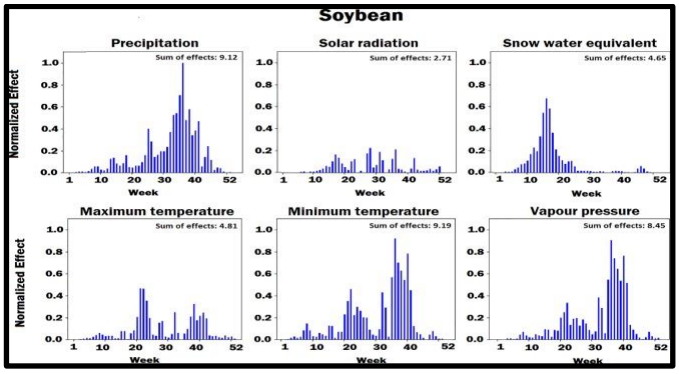

Fig 5.1. Bar plot of estimated effects of six weather components on corn measured for 52 weeks of each year, starting from January. The vertical axes were normalized across all weather components to make the effects comparable.

2. SVM Algorithm:

For binary classification problems, SVM (Support Vector Machine) is a supervised machine learning approach. Picture 6.2 Maximum Margin and Hyperplanes The algorithm's main objective is to find the plane with the highest margin, or the greatest separation between the data points of the features being plotted, albeit any number of hyper planes may be drawn. The classification will be more accurate the more away it is. The data points are far from all the other points from the optimal hyperplane, as illustrated in fig. 6.2, resulting in a maximum margin. Data collection, data preparation, feature selection, SVM model training, model validation, and prediction are a few of the steps that can be included in SVM.

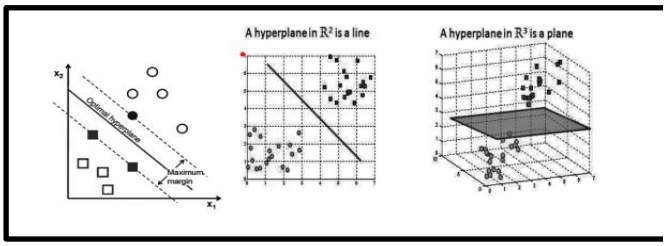

Fig 5.2: Maximum Margin and Hyper planes

6. UML DIAGRAMS

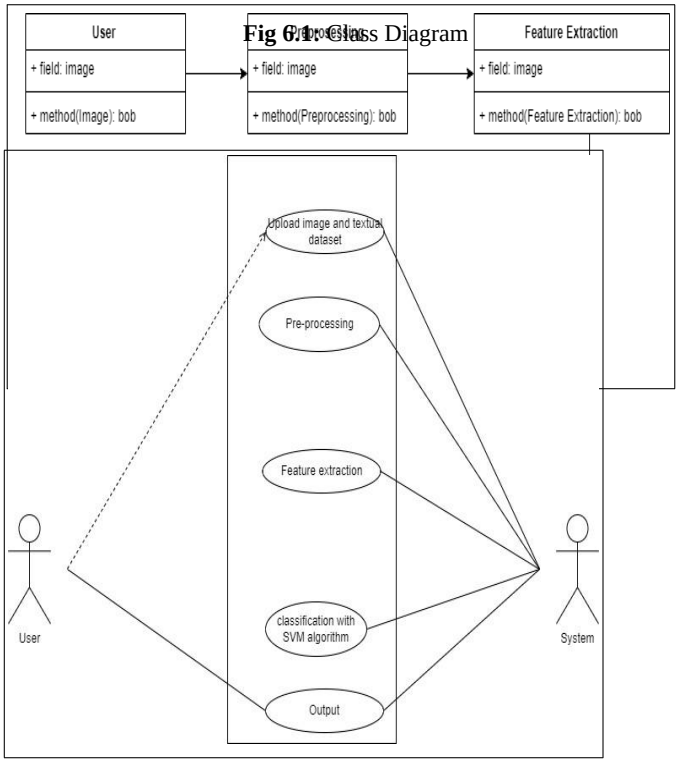

Fig 6.1: Class Diagram

Fig 6.2: Use case Diagram

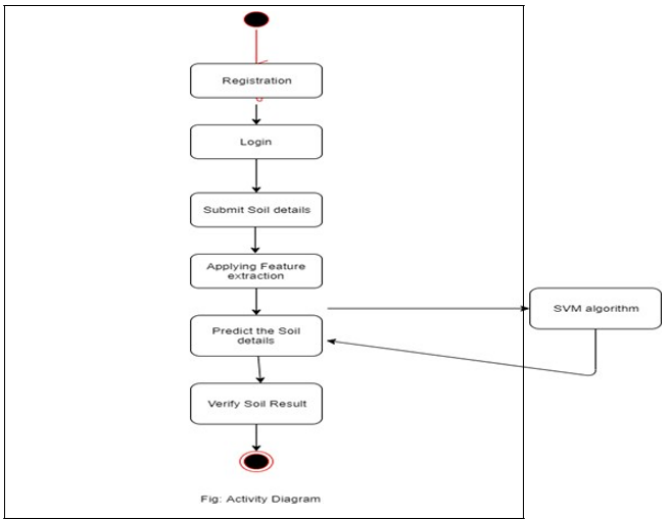

Fig 6.3: Activity Diagram

7. SYSTEM SPECIFICATION

Hardware Specifications
- System: Intel I5 Processor and over.
- Hard Disk: 150 GB
- Ram: 8 GB

Software Specifications
- Operating system Windows 7 or further.
- Coding Language: Python
- IDE: PyCharm.

8. EXPECTED RESULT

The expected outcome will rely on a number of variables, thus in order to get the best results, it is crucial to carefully prepare the dataset, choose the right features, and optimize the model's hyperparameters.

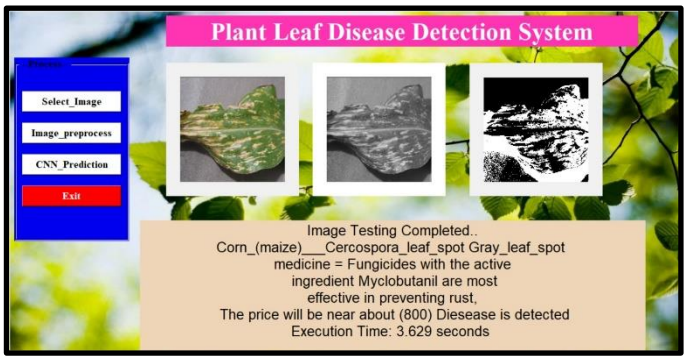

Fig 8.1. User Interface

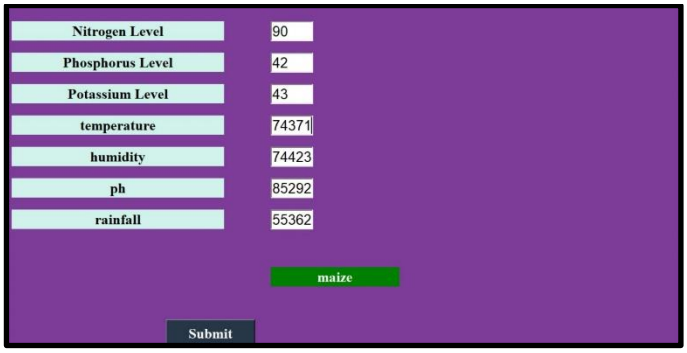

Fig 8.2: Results

9. CONCLUSION AND FUTURE WORK

Agriculture contributes to the economy of our country. However, this falls short of using modern machine learning techniques. Therefore, our farmers should be familiar with all the most recent machine learning technology and other cutting-edge techniques. These techniques help to increase agricultural productivity. In order to boost crop output rates, a variety of machine learning approaches are applied in agriculture. These techniques can help with agricultural problems. We may also assess the accuracy of the yield by looking at several ways. Therefore, we can improve performance by contrasting the accuracy of distinct crops.

10. REFERENCES

[1] https://en.wikipedia.org/wiki/Agriculture

[2] https://en.wikipedia.org/wiki/Data*analysis JeetendraShenoy,YogeshPingle,\IOTinagriculture*", 2016*IEEE*.

[3] M.R. Bendre, R.C. Thool,, V.R. Thool, "Big Data in Precision agriculture", Sept, 2015 NGCT..

[4] Monali Paul, Santosh K. Vishwakarma, Ashok Verma, "Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining approach", 2015 (2002) The IEEE website. [Online]. Available: http://www.ieee.org/

[5] International Conference on Computational Intelligence and Communication Net- works. *FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.

[6] N. Heemageetha, "A survey on Application of Data Mining Techniques to Analyze the soil for agricultural purpose", 2016IEEE.

[7] https://en.wikipedia.org/wiki/Linearregression*Y.Jee vanNagendraKumar, Dr.T.V.RajiniKanth,\GISMAPBasedSpatialAnalysis o f Rain f allDatao* 2088 − 8708

[8] B Sankara Babu, A Suneetha, G Charles Babu,Y.Jeevan Nagendra Kumar, G Karuna, " Medical Disease Prediction using Grey Wolf optimization and Auto Encoder based Recurrent

[9] Neural Network", Periodicals of Engineering and Natural Sciences, June 2018 ISSN 2303-4521 Vol.6, No.1, pp. 229 240