[1]Sathis Kumar K
Arulanandam K[2*]

Hybrid Machine Learning based Crop Prediction Model

Journal of Electrical Systems

Abstract: - The manufacturing of agricultural goods has always been an essential component of economic growth, and it has had a significant influence on the economic well-being of our country. In addition, as a result of the quick development of science, the agricultural sector has emerged as one of the most important areas in which to address challenges associated with farming. These challenges include land, the flow of groundwater, catastrophic events, herbicides, and pesticides. It is essential to the growth of harvests in farming that an adequate quantity of precipitation be observed throughout all stages of yield development. Along with crop prediction, this paper provides clear information about the quantities of soil ingredients required and their associated costs by applying deep learning techniques to the suggested model. When compared to the current approach, it improves precision. By analyzing the data provided, it aids producers in making informed business decisions. The land's climate and dirt are factored in to estimate a reasonable harvest. The goal is to introduce a Python-based system that makes use of strategic thinking to foresee the most productive harvest under specific circumstances while minimizing associated costs. Machine Learning is handled by the Support Vector Machine (SVM) algorithm, while Deep Learning is represented by the LSTM and RNN algorithms.

Keywords: Machine Learning, Support Vector Machine, Deep Learning, Long short-term memory, Recurrent Neural Network

1. Introduction

In the context of the food supply chain, the management of food security has evolved into a very challenging and important problem. The ability to accurately forecast crop output reveals numerous important opportunities for improving the administration of food safety along the supply chain of food. The prediction of crop output provides information that can serve as the foundation for many significant decisions relating to food security, such as the development of policies and the trading of commodities.

On the other hand, estimating the crop yield is not a simple task because it is influenced by a large number of controllable and uncontrollable factors. Some of these controllable factors include the amount of irrigation, pesticide, and fertilizer that is applied, while other factors include the market, subsidies, and the weather. There were a number of ways to predict the harvest, but the vast majority relied on the farmers' years of experience and the average yield from previous years. While some scientists agree with Schlenker and Roberts that annual crop output fluctuates and does not follow a linear pattern, others disagree. Models of crop yield production informed by data are given as the gold standard in the scientific literature. Not only are data-driven models for crop yield prediction cheap and easy to apply, but they also improve the precision of data gathering mechanisms.

India's economy relies heavily on agriculture. The weather plays a huge role in India's farming output. Rainfall is crucial to the success of rice farms. Helping farmers optimize crop yields requires accurate predictions of future crop productivity based on careful analysis. Predicting crop yields is a critical issue in agriculture. In the past, farmers would extrapolate their output from the year before. Consequently, we can forecast crop yield with the aid of these algorithms, which were developed specifically for this kind of data analytics in crop prediction. It employs a random forest method. As a result of these numerous innovations, many people now devote their time and energy to the production of artificial goods, many of which are hybrids that contribute to an undesirable way of living. People in the contemporary era lack an understanding of when and where to plant crops for optimal growth. Seasonal climatic conditions are also being altered by these cultivating methods,

[1] Research scholar, GTM College, Gudiyattam, Affiliated to Thiruvalluvar University, Serkkadu, Vellore, Tamilnadu, India – 632115.

[2]Head & Department of Computer Applications, GTM College, Gudiyattam, Affiliated to Thiruvalluvar University, Serkkadu, Vellore, Tamilnadu, India – 632115.

Email : sathish739@gmail.com, arulanandam@gtmc.edu.in*

which works against fundamental assets like soil, water, and air and ultimately leads to food insecurity. After looking into things like weather, temperature, and a number of other variables, we haven't been able to come up with a viable solution or technological advancement to help us out of this mess. Agricultural expansion in India can be boosted in a number of ways. Multiple strategies exist for maximizing harvests and enhancing produce quality.

When it comes to making decisions concerning agricultural risk management, having an accurate prediction about the past history of the crop yield is an essential thing that can be used. Before beginning cultivation on the field, the farmer will first determine the potential yield of the produce. This provides producers with some insight into what may or may not occur with regard to a specific crop during that year. If the circumstances are favorable, this could work to their advantage and help them achieve higher crop productivity. It may also assist them in lowering losses brought on by circumstances that are unsuitable. The majority of researchers in this field are currently employing machine learning techniques in order to model and validate the difficulties associated with precise agriculture. The use of machine learning is a significant move toward ensuring a prosperous future for farming. The grower can now sit at the center of choices that are predictively informed by the abundance of technology that can help mitigate risk and improve sustainability.

The driving force behind all of these technologies is machine learning. The use of machine learning has allowed for the creation of personalized dashboards for farmers to access all of their crop data, as well as problems relating to their crops and yield, and effective solutions for managing the problem. In addition, the majority of contemporary agricultural practices involve the utilization of robots that are purpose-built and pre-programmed to perform a variety of tasks associated with agriculture. Their devices assist them in gathering data, conducting an analysis of the issues, and putting a solution into action for the issues. These machines are very helpful to farmers, and they also complete their tasks much more quickly than human laborers.

Machine learning (ML) methods can be found in many different industries, from grocery shops that analyze customer behavior to telecommunications firms that foresee how customers will interact with their devices. Agribusiness is another area where machine learning has been used for some time. Many models have been suggested and validated for the prediction of crop yield, which is one of the difficult problems that precision agriculture aims to solve. This problem calls for the integration of numerous data sets due to the complexity posed by the interplay of so many factors affecting crop yield, such as climate, weather, soil, fertilizer application, and seed type. As a result, it appears that estimating crop output is not a straightforward process, but rather one that involves several intricate steps. These days, crop yield prediction models are able to make a reasonable estimate of the actual yield, but it would be preferable to have an even better performance in yield projection.

Machine learning, a branch of AI focused on instruction, is a technique that, when put into practice, can enhance output prediction on the basis of a variety of factors. The concept of machine learning (ML) refers to the process of extracting patterns, correlations, and information from datasets. It is necessary for the models to be educated with the help of datasets, in which the results are represented based on previous experience. The predictive model is constructed by using a number of different features; consequently, during the training phase, the parameters of the models are established by making use of historical data. Part of the historical data that was not used for training is used to evaluate the system's efficiency during the testing process.

Depending on the research challenge and research questions, a machine learning model might be descriptive or it might be predictive. The purpose of descriptive models is to acquire knowledge from the data that has been collected and to explain what has happened. The goal of any predictive algorithm is, obviously, to foretell the future. There are many challenges in ML study that must be met before a high-performance predictive model can be built. In addition to making sure the algorithms and platforms they're built on can handle the volume of data being processed, it's also important to make sure you pick the right algorithms for the issue at hand.

Over the past decade, machine learning has been the subject of intensive research, and it is now being used to improve crop forecasting around the world. The value of county-level crop production prediction models has been proven in a number of studies. However, there has been a dearth of research on forecasting yields on a farm's size. Since there is not enough money for sustainable agriculture and the costly collection of satellite images, there is not enough ground-truth data for output on a farm-scale. However, these obstacles appear to be thinning out in the agricultural industry. Since 2017, the Norwegian government has released comprehensive reports on the country's agriculture, including data on a national scale.

1.1 Crop Yields and its importance

The amount of edible seeds or grains harvested from a specific plot of ground is known as its crop yield. It is typically given in terms of pounds per acre or kilograms per hectare. A farmer's productivity in a field over a particular time period can be evaluated, in part, by looking at the average crop yield per acre. It represents the culmination of farmers' time and energy spent cultivating crops in their farms, making it a crucial indicator of their success as business owners. Given this, it's no surprise that "how to increase the average crop yield per acre" is a perennial question for most farmers.

Farmers have been thinking about and working on the problem of how to increase crop yields for millennia. The effectiveness of the solutions discovered varied widely. Not only can today's cultivators learn from the wisdom of their ancestors, but they can also leverage the advances in science and technology that have been made in recent decades. Let's take look at the most common techniques used to boost agricultural output and discover how combining traditional wisdom with cutting-edge technology can help producers succeed.

The following are the procedures to increase the yielding of crops

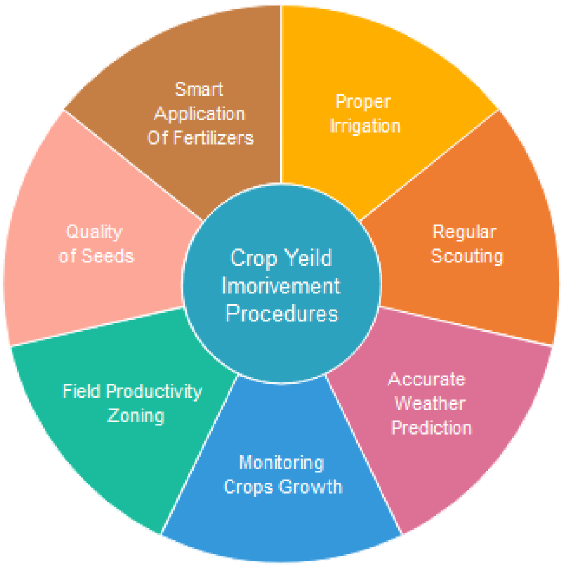

Fig1: Procedures to increase the yielding of crops

- ***Quality of Seeds***

The agricultural output is directly proportional to the quality of the seeds used by producers to plant their crops. It follows that only certified seeds, which have been proven to satisfy all relevant quality controls, should be sown on agricultural land if farmers are to maximize their crop yields. The price of certified seeds may be higher than the price of seeds that are not certified, but the investment will be worthwhile in the long run because the quality of the seeds is one of the most important variables that determine the amount of food that can be harvested from a given area. In addition to that, one of the environmentally responsible ways to improve crop yield is to ensure that you only plant seeds of a high quality. In the event that it becomes essential, a farmer can verify the viability of specific seeds by contacting an appropriate seed company and having them conduct special tests on a designated plot of land.

Aside from this, it is essential to keep in mind that the quality of seeds is not something that is unchangeable and unaffected by damage over time. The instant seed grains are sown into the ground, they require protection from the elements. Coating the seeds is one method of treating seeds that can be used for this purpose. To improve the characteristics (weight, size) and/or provide some active compounds that are called to protect seed grains from plant diseases and boost their growth, seed coating is the process. The goal of this process is to improve the characteristics of the seed grains.

- ***Field Productivity Zoning***

Before planting, it is essential for a farmer to have a good understanding of the productivity of the field that is going to be planted in and, if necessary, to identify particular areas in the field in which plants tend to develop better. Productivity zoning is the name given to this procedure. Growers can use this method to plant seeds more densely in regions that have a higher potential yield, increasing the likelihood of increased crop output, while minimizing the amount of seed that is lost in regions that have a lower potential yield. In addition, by dividing their fields into distinct zones, farmers can better care for low-yield regions and take whatever measures are required to improve the soil's fertility and other crucial factors for growing a particular crop.

The modern tools at a farmer's disposal make the process of productivity zoning their areas faster and more precise than ever before. The EOSDA Crop Monitoring software is used by a significant number of farmers for this specific purpose. Growers are able to generate field productivity maps based on historical data thanks to the company's high-precision technology and satellite-driven data. As a result, they are able to clearly identify the regions of a particular farmland that have the highest and the lowest levels of productivity.

- ***Monitoring Crops Growth***

In order to quickly identify any problem that can emerge on a given farmland (whether it be related to pest infestations, plant diseases, weeds, etc.) and that can affect the crop output, regular plant health monitoring is essential from the earliest stage of development through budding and all the way up to harvesting. This monitoring should continue until the plant is harvested.

Regular satellite monitoring of land plots, for example, allows farmers to easily track the growth state of plants and carry out crop yield estimation using remote sensing. The development of the health status of plants provides important information that can be used to decide on potential interventions to meet the requirements of crops.

Despite the fact that satellites do not directly measure the stages of plant development, spectral indices can provide useful information. In addition, EOSDA Crop Monitoring provides statistics on daily temperatures, analyzes how those temperatures change over time, and computes the sum of all of those temperatures. Because of this information, software is now able to determine the stages of plant development for a wide range of plant species and represent the correlations between those stages and other data, allowing farmers to make decisions that are informed by sufficient context.

- ***Accurate Weather Prediction***

The typical crop yield that can be expected from an acre of land in a particular field is greatly influenced by the weather. The climate conditions have a predominate influence on the development of plants and, as a result, on yields, even if the soil quality and the variety of seeds planted are the same. This is particularly true when taking into consideration the effects that climate change has on agriculture in various regions and in general across the planet. Farmers have the opportunity to consult the most cutting-edge technological solutions in agriculture, which can assist them in obtaining accurate weather prediction, so that they can work effectively and cooperatively with the weather, which is a significant but uncontrollable component.

For instance, farmers who employ tools like EOSDA Crop Monitoring software in precision agriculture have a better chance of protecting their harvest from the adverse effects of weather fluctuations. This is an extremely beneficial feature. Growers can benefit from high-precision weather forecasts by determining which agricultural techniques they are able to carry out. Given all of this information, it is important to highlight the fact that precise weather forecasting, which is a component of precision farming, can make a significant contribution to increasing agricultural yields on a specific plot of farmland.

- ***Regular Scouting***

The correct growth of plants and an increase in crop yields can both be helped along by scouting, which is an essential component of the management of agricultural businesses. When a grower has large fields, it is often difficult and inefficient to survey the land one acre at a time because doing so requires a significant investment of both time and resources.

This is an opportunity that EOSDA Crop Monitoring seizes and uses to its advantage by providing convenient scouting assignment management. The time-saving Crop Scouting solution is designed to take the place of the laborious work performed by human specialists in fields. All that is needed of you is to log into your Crop Scouting app, look for problem areas on your land plot that were discovered by a satellite, mark these areas on the map, and set up closer scouting of these zones.

The next step is for the scouts to investigate the problem at the site, during which they will snap pictures, add them to the app, and fill in any relevant information. In this way, a farmer can set up numerous scouting duties simultaneously, all of which can be easily managed and monitored as they move forward toward completion. A field scouting solution that is both time-efficient and highly accurate, like this one, helps farmers in their day-to-day work and has the potential to help them increase their harvests.

- ***Proper Irrigation***

A reliable irrigation system is essential for producers who wish to increase their average crop yield per acre. The development of plants and, as a result, the amount of food that can be harvested from those plants is directly impacted by the amount of water that is made available to those plants. The accuracy of weather predictions is essential to the efficiency of agricultural irrigation. Access to weather forecasting at a hyper-local level is now possible, thanks to modern technologies such as specialized applications and software designed specifically for farming. It paves the way for precision irrigation and gives farmers the ability to organize the irrigation of their fields in the most precise and effective fashion possible by planning ahead and organizing it.

- ***Smart Application Of Fertilizers***

Fertilizers should be used moderately and sensibly despite their beneficial effects on soil health, plant development, and crop production. The soil quality and agricultural output are both adversely impacted by excessive fertilizer use. Different zones of a single field may have varying requirements for soil fertilization, so it is best to apply fertilizers carefully in response to specific needs. This method of precise fertilization of fields maintains healthy soil, which in turn increases the typical output per acre.

EOSDA Crop Monitoring software is one example of a technological option that could be used here. The satellite-based field zoning function allows farmers to see which parts of their farms need more attention than others. Precision farming is widely used by farmers because it is one of the most environmentally responsible ways to boost crop yield.

1.2 Crop Prediction Methods

In order to maximize harvests, farmers must tend to their plants from the beginning of the growth season until harvest time. Farmers use a variety of chemicals and other substances to combat weeds, pests, and diseases that can reduce crop output. Pesticides, fungicides, plant growth enhancers, adjuvants, desiccants, herbicides, and other similar chemicals are common examples.

In general, there is a wide range of strategies for safeguarding crops. Controlling invasive species and preventing plant diseases are two of the most important.

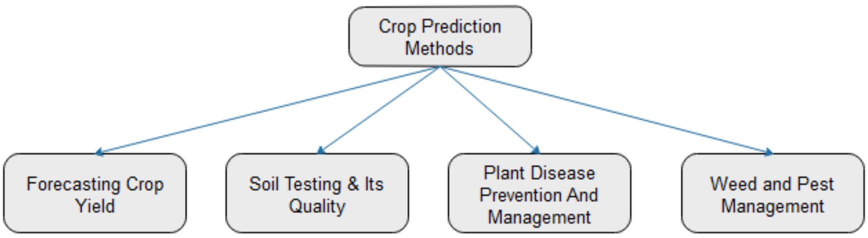

Fig2: Crop Prediction Methods

- **Weed and Pest Management**

Weed and pest management pose the biggest challenges to farmers during the growing season. One weed, for example, can generate more than 10 million weed seeds, which, if not eradicated in a timely way, can significantly reduce yields and cause problems for years to come. In order to effectively combat pest infestations, farmers must adopt a systemic approach to pest control. Due to their ability to rapidly adjust and reproduce, pests pose a serious threat to the crop yield on any given piece of farmland, necessitating rapid action from the farmers who manage it.

- ***Plant Disease Prevention And Management***

Plant illnesses are another major factor that can reduce a farmer's harvest. Farmers can use a number of different strategies for preventing and controlling plant diseases, such as choosing disease-resistant or disease-tolerant plant types, treating seeds with fungicides, spraying developing plants with pesticides, and so on. For best results, farmers should implement these or similar strategies at the optimal moment. Remember that timely plant protection is critical for the indicator of crop yield per acre, as problems are simpler to fix and fewer field acres are lost if they are caught early.

- ***Soil Testing & Its Quality***

One of the most influential aspects of harvest success is the soil's condition, or fertility. The price a farmer pays to cultivate a particular plant depends on a number of factors, including the soil quality, which affects agricultural output because different plants have different nutrient and water requirements. Farmers who want higher crop yields need to pay careful attention to the state of the soil on their fields. Crop succession is one of the most effective methods for preserving soil health. Plant rotation improves agricultural production by reducing soil depletion and interrupting pest life cycles, leading to a higher average crop yield per acre.

- ***Forecasting Crop Yield***

Predicting crop yields is crucial to meeting the world's dietary needs. Analytical data on projected crop yields is used by governments around the globe to make educated decisions about their import and export policies. Seed-breeding firms need to predict how newly developed hybrids will fare in a variety of environmental and land settings before committing resources to further developing those sets of seeds. The ability to accurately forecast crop yields is used by farmers to make informed agribusiness choices.

Predicting crop output using remote sensing satellite data is one of the newest developments in industrial agriculture, and it has already been proven effective in various parts of the world. The reliability of remote sensing estimates of agricultural yield is, however, sensitive to a wide range of variables, including climate (especially extreme weather), soil, pests, and disease. Agricultural risk management and crop yield prediction also rely heavily on accurate historical records of yields in a given area.

Many different machine learning methods are currently in use around the globe to forecast agricultural output. Large amounts of data are needed for crop yield prediction software, but it may not always be effective due to factors like weather changes.

2. Related Works

The manufacturing of agricultural goods has always been an essential component of economic growth, and it has had a significant influence on the economic well-being of our country. In addition, as a result of the quick development of science, the agricultural sector has emerged as one of the most important areas in which to address challenges associated with farming. These challenges include land, the flow of groundwater, catastrophic events, herbicides, and pesticides. It is essential to the growth of harvests in farming that an adequate quantity of precipitation be observed throughout all stages of yield development. Understanding this occurrence is crucial for farmers because it can help them estimate how much water is needed for agriculture that requires irrigation. Even the best rain seasons won't always be enough to boost crop yields. It is a challenging endeavor that requires careful and dependable consideration of a large number of variables, such as relative humidity and temperature, in order to accurately forecast the possible amount of precipitation that will fall as well as the produce and the amount of water that will be needed. In previous times, the viability of the monsoon was estimated by taking into account three crucial aspects: the level of humidity, the temperature, and the amount of precipitation. Investigate a sizeable percentage of precipitation heaviness while analyzing a number of different research mechanisms that have been used throughout history. M. E. Patil et al., (2022) use a hybrid strategy that predicts crop production based on annual rainfall using a combination of logistic regression and random forest (LRRF).

In India, agriculture is an industry that is expanding at a rapid rate. The most important problems facing the agricultural industry today include crop yield prediction, the level of soil nutrients, intelligent irrigation systems, crop monitoring, and other problems. This research by A. K. Sharma and coworkers (2022) examines the smart farming system that relies on the Internet of Things. The primary objective of this research is to improve the accuracy of the current Crop Yield Prediction while simultaneously reducing the amount of human involvement and moving closer toward fully automating the system. The Crop Yield Prediction is responsible for controlling and monitoring the crop yield amount. The camera's sensor updates the visible value and stores the new information in the cloud whenever the Crop Yield level rises. The technique we propose for predicting crop yields makes use of a Hybrid Deep Learning Algorithm for Smart Agriculture.

Due to the fact that it is dependent on a wide variety of variable parameters, such as environmental, weather, soil, and climate factors, crop yield prediction is currently one of the most cutting-edge tasks that is also one of the most fascinating and difficult to complete. In recent years, machine learning has emerged as one of the most useful methods for estimating crop yield. This article by A.P.S. Manideep (2022) provides a machine learning framework for predicting crop yields by using data on both the crops and the weather. It uses past data to compare the performance of eight machine learning techniques for making predictions about crop yields in various regions of India between 2001 and 2016. These techniques include regression, decision trees, random forest, support vector machine, and gradient boosting. In addition, the root mean square error (RMSE) was calculated to be 9433.7 for the dataset when using the random forest method. This was determined by analyzing the findings.

Predicting the output of a crop is one of the more difficult tasks involved in the agricultural industry. Artificial Neural Network (ANN), a machine learning algorithm, and Multiple Linear Regression, a statistical model, have both been the subject of extensive research in the agricultural sector in the hopes of improving the accuracy of crop output forecasts. The article "The Intrinsic Relationship between MLR and ANN" by Maya Gopal, P.S. (2020) investigates this relationship. Within the scope of this research project, an MLR-ANN hybrid model has been suggested as a method for accurately predicting crop yields. The accuracy of a prediction is measured using a model of the suggested hybrid model, in which the MLR intercept and coefficients are used to set the weights and bias of the ANN's input layer. To make a reliable forecast of paddy crop yield, a synthetic feed-forward neural network (ANN) was created and taught with back propagation. The weights and bias of the input and hidden layers in a conventional ANN model are arbitrary when the model is first initialized. Instead of using random values for the weights and bias during the initialization process of this composite MLR-ANN model, the input layer weights and bias are initialized by using the MLR's coefficients and bias. Utilizing performance measures, a comparison is made between the accuracy of prediction made by the hybrid model and that of the ANN, MLR, Support Vector Regression (SVR), k-Nearest Neighbour (KNN), and Random Forest (RF) models. The amount of time required for calculation for hybrid MLR-ANN as well as conventional ANN was determined. According to the findings, the hybrid MLR-ANN model that was suggested provides a higher level of accuracy compared to the conventional models.

Taking into consideration the current predicament that Indian producers are in. We have also seen that there have been a significant number of people who have taken their own lives in India over the past several years. The reason for this is because the weather patterns in our country are constantly shifting, and our government is in a state of constant flux. There are times when farmers are not conscious of the types of crops that will thrive in their particular soil. Soil quality can be determined from factors like temperature, moisture, and pH value, and the purpose of the Ishwarya R et al., 2022 initiative is to help farmers interpret these factors. In addition to determining which crops will do best in a given area, this initiative estimates how much those crops will produce before farmers even plant them. Machine learning is the most effective method for resolving this issue because it can be used to implement a wide range of algorithms that are both accurate and precise in their predictions of crop yield and useful in advising farmers on the best crop to grow and the amount of fertilizer to use based on the quality of their soil. Machine learning using these algorithms is the most efficient way to solve this issue in the real world.

Farming is the primary occupation that is primarily regarded as a culture rather than a job, and farming is the linchpin of our economy due to the fact that farming is the means that brought forth human advancement. India is a nation that demonstrates a greater interest in farming and also produces all different kinds of crops; the profitability of its harvest is a primary driver of India's economy. As a consequence, we are able to state that agriculture is a primary pillar of support for all of our nation's businesses. Because every state in India is responsible for a different kind of agriculture due to the wide range of local climates, the selection of each produce is an important part of the overall process. The value of the product, the price set by the government, the conditions of the weather, and the price set by buyers in the private market will all play a role in determining which crop is selected. Agriculture is one area that desperately needs numerous advancements in order to improve its contribution to the Indian economy. Implementing AI mechanisms that can be used in the same way and to the same effect across a variety of cultivating areas can help us enhance agriculture. Now that there have been so many developments in the area of machines and their improvements, we can use them to cultivate the valuable and detailed data concerning a variety of problems, which plays a crucial role in the process. In this article, Mohammed Ali Shaik (2022) provides us with a framework for thinking about how to solve a wide range of issues related to agriculture and agricultural fields by combining ambitious techniques with a harvest-based strategy. This makes it easier for producers to select the most profitable crop, which in turn increases their income and contributes to the overall economic growth of our country.

In order to extrapolate city yields from weather data, Keach Murakami (2021) created a machine learning model. This research looked at how the weather in Hokkaido, Japan, affects the harvest of winter wheat. This island has a wetter climate than most wheat-producing areas because of the high annual precipitation average and the large amount of snowmelt water that flows off of the mountains in the spring. These findings corroborated our knowledge of the weather's effects on wheat production, suggesting that explainable machine learning could aid in the meteorological prediction of crop output in humid climates.

The significance of agriculture is not confined to our day-to-day lives; rather, it is an important sector that significantly contributes to the economic development of any nation. As a result, one of the most important steps in the process of obtaining competitive crops is to improve the quality of the agricultural yields by utilizing modern technologies. Data mining is a relatively new research topic that is becoming increasingly important in the agricultural industry, particularly in the forecasting and evaluation of crop yield. In order to predict the impact of different factors on crop yield quality, Rashid Ismael, H. (2021) examines the use of various data mining classification algorithms. Location, time of year, and output are all examples of such factors. Each algorithm, from the K-nearest neighbour to the naive Bayes to the random forest to the support vector machine, has its performance measured and contrasted. The contrast takes into account the error rates and precision measurements. In terms of precision, the SVM algorithm performed best (76.82%) while the KNN algorithm performed worst (35.76%). The KNN algorithm had the highest error number, which were 111.8855. Additionally, the projection assisted farmers in raising their incomes and making them more stable.

The prediction of farm-scale crop yields is a logical progression of sustainable agriculture, which aims to produce a substantial quantity of food without depleting environmental resources or polluting the environment. The scope of recent research on crop yield generation has been confined to predictions on a regional scale. Regional crop yield predictions often run into snags when attempting to account for differences in local yields resulting from differences in farm management decisions and the state of the field. We found that it would be helpful to build a large, reusable dataset of farm-scale agricultural yield production data in order to accomplish our ground-truth prediction objectives at the farm scale with high accuracy. We use Sentinel-2 satellite imagery, weather records, agricultural records, grain transportation records, and cadastral data, to name a few examples of data covering multiple time periods. To facilitate learning across numerous time periods, Engen et al. present a deep mixed neural network model. (2021).

3. Proposed Methodology

3.1 Data Collections

The present research concentrated on farms situated in a variety of locations across the Asian continent, totaling approximately 183,285 km2 and accounting for 1% of Iran's total land area. In order to accomplish what the research set out to do, two distinct kinds of data were gathered: 1) the production of agricultural goods, and 2) intelligence regarding the weather. Collecting information about product types, water usage, and harvest totals are just some of the uses for the Spriter-GIS system. Two meteorological stations were set up on the location for the purpose of data collection, including measurements of rainfall, solar radiation, and temperatures. On the other hand, this research focused on agricultural goods like wheat, barley, potato, and sugar beets because they are the most important crops produced in this area. Table 1 provides a rundown of the information that was gathered in its entirety. The model is fed information from a massive dataset of generic crop information that includes agricultural characteristics. The features come from an independent data source. The databases were compiled using tools found on the website kaggle.com. The entire size of the harvest archive is 7841 kilobytes. Prediction factors in this set include not only temperature and precipitation but also pH, relative humidity, and area. Wheat, rice, corn, sugarcane, and a host of other commodities are represented in this dataset. There is a spectrum of numbers within which each prediction parameter can be utilized for any given crop. When wheat is selected as the crop, for instance, any of the numbers in the dataset can be used to populate the wheat prediction parameters. These values can be found in the dataset. It is the same for each and every crop in the dataset that is accessible.

Table1: The Testing and Training Dataset Distribution

Crop species	Total number of samples	Samples in the training period	Samples in the testing period	Testing percentage (%)
Wheat	651	582	69	10.61
Corn	91	48	43	47.25
Potato	231	156	75	32.46
Sugar Beet	189	128	61	32.27
Paddy	658	437	221	33.58

3.2 Proposed Frame work

The techniques of machine learning and deep learning are put into action within the framework that has been suggested in order to make accurate forecasts regarding crop production. The suggested model conducts an experiment using a dataset consisting of crops. As the climatic and soil parameters are taken into consideration, the crop is selected based on the present atmosphere, the soil, and the constituents of the soil. Deep learning is used to accomplish a large number of successful calculations, and it is also used to determine which crop is the most suitable choice when there are multiple possibilities. Accurate forecasts of crop yields can be generated with the help of this method. As shown in figure 3, the SVM algorithm is performed using the machine learning technique, whereas the LSTM and RNN algorithms are carried out using the deep learning methodology.

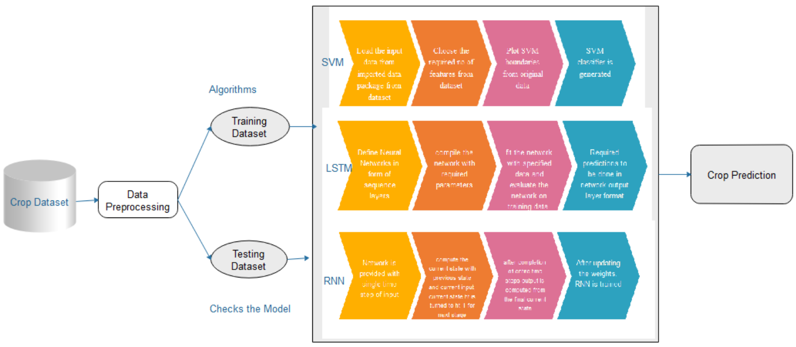

Figure 3 The suggested model's structure for making crop predictions

Implementation steps

- Step 1: The crop dataset, which includes a number of characteristics, should be loaded.
- Step 2: Bring up the helpful libraries and software programs.
- Step 3: The initial processing of data is carried out.
- Step 4: Separating the data into a training set and a testing set is the first step in getting ready to analyze the dataset.
- Step 5: Following this step, machine learning (SVM algorithm) and deep learning (LSTM, RNN) techniques are applied in order to construct a model. This model then makes a prediction regarding the most optimal crop that should be cultivated.
- Step 6: The performance of the model is evaluated using the test collection. The model will generate an error message that reads "value mismatch or wrong prediction" if any garbage value is provided as input.

Algorithms Used for proposed framework

SVM (Support Vector Machine)

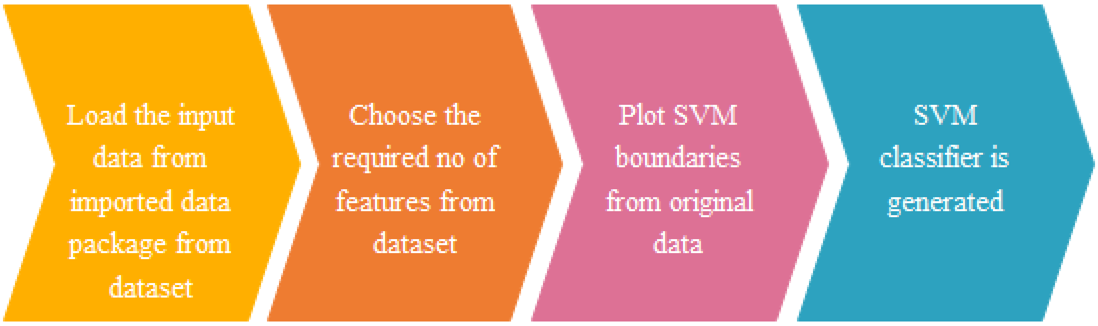

Fig4: SVM Flowchart

- Bring in the required software programs.
- Load input info.
- Select the necessary amount of features from the collection..
- By referring back to the initial data, plot the SVM confidence intervals.
- Establish a number to serve as the regularization parameter.
- At long last, the object that will be used by the SVM classification is generated.

Long-Short Term Memory (LSTM)

- Define a neural network in Keras using a succession of layers as its building blocks.
- Compile the network, which requires many different parameters to be described.
- A training network requires two sets of data: input patterns (represented by a matrix) X and target patterns (represented by an array) y. Conform the network to the training set.
- Use the training data to perform a study on the network. Test and validation datasets are used to assess the model's performance, but this is rarely done.
- Perform the necessary predictions, which must be accomplished in the manner specified by the output layer of the network.

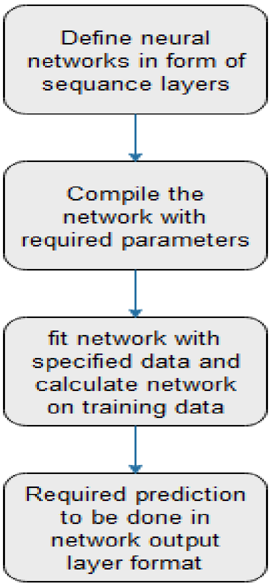

Fig5: LSTM Flowchart

Recurrent Neural Network (RNN)

- The input to the network consists of a single time increment.
- Create the current state by computing it with the assistance of the state before it and the input that is currently being used.
- The present state, denoted by ht, will be denoted by ht-1 for the next time step.
- It is possible to take any amount of time steps necessary, depending on the issue. All of the information from the preceding states has been combined into one.
- After all of the time steps have been executed to completion, the ultimate present state is incorporated into the calculation of the output.
- Back-propagation of the error towards the network is what's needed in order to get the weights updated. As a result, the RNN is educated.

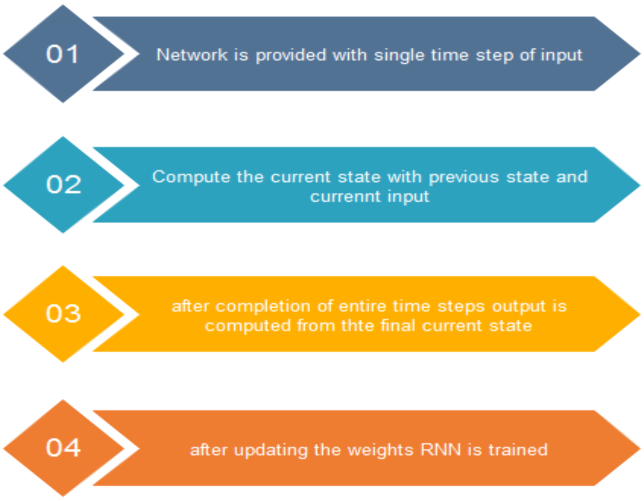

Fig6: RNN Flow Chart

4 Implementation and Result

The first stage in putting the planned research into action is loading the gathered dataset of crops. The subsequent procedure is data preprocessing, which follows the import of necessary libraries and packages. Separate streams of learned data and test data are spit out. Eventually, the required AI programs will be built into a model. Consequently, this model will supply the crop that is best suited to be grown on a given plot of ground. An agricultural dataset is used for the analysis, and it was found on the website kaggle.com. Wheat, rice, sugarcane, and a plethora of other products are all included. A few projection factors, such as pH value, temperature, precipitation, relative humidity, and area, are included. The Trained set and Test set are required by machine learning and deep learning techniques to build a predictive model. Predicted outcomes are based on historical occurrences that have been surveyed for training purposes. Currently available data is the Test results from the survey. The output of the code is a graphical representation of the results for each of the input values. These factors include not only the amount of water and sunlight available, but also the amount of fertilizers and herbicides used. Based on these factors, a yield estimate can be made:

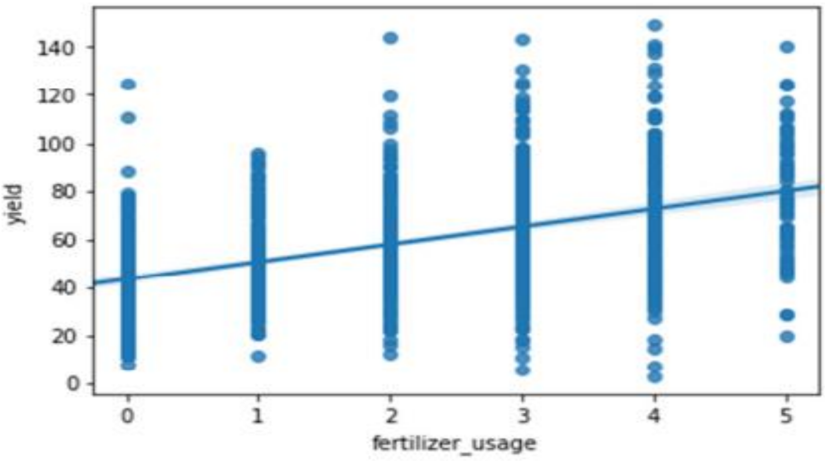

Fig 7. Production versus Fertilizer Input

The above graphical representation shows the yielding of crops according to usage of fertilizers in the agriculture fields. These graphs are drawn using the SVM classifiers. In a scatter plot, the x and y numbers are used to determine where each data point will be placed. By using above figure 7, the crop yield is calculated with respect to usage of fertilisers.

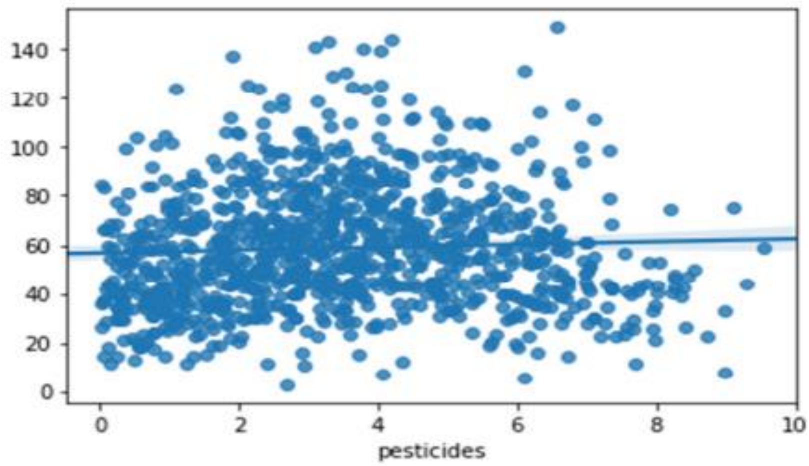

Fig 8. Ratio of Crop Production to Chemical Pest Control

The above graphical representation shows the yielding of crops according to usage of chemical pest in the agriculture fields. These graphs are drawn using the SVM classifiers. In a scatter plot, the x and y numbers are used to determine where each data point will be placed. By using above figure 8, the crop yield is calculated with respect to usage of chemical pest.

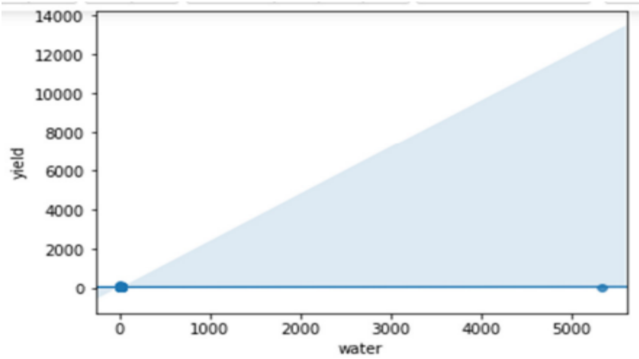

Fig 9. Produce versus Use of Water

The above figure 9 shows how the yielding of crops is increased by increasing the usage of water. The above graph is drawn using the svm classifier. In a scatter plot, the x and y numbers are used to determine where each data point will be placed. From the above figure crop yield also calculated in this case.

The below figure 10 represents the yielding of crop in the unit area of land. By using the SVM classifier this graph is drawn and also calculated the yielding percentage as per the unit land. In a scatter plot, the x and y numbers are used to determine where each data point will be placed.

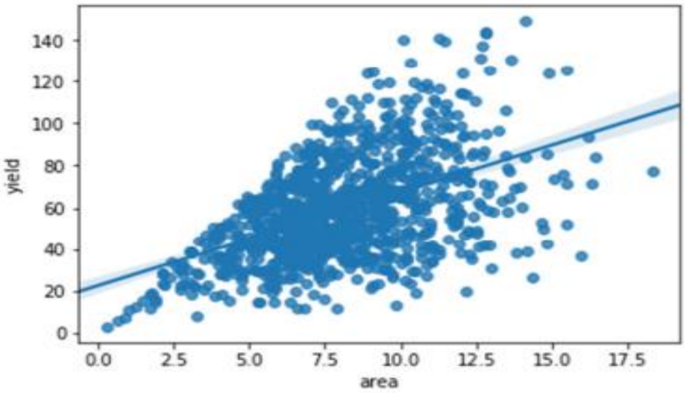

Fig 10. Productivity per Unit of Land

The yield versus uv is depicted in the aforementioned diagrams. These charts are generated with the help of the SVM algorithm. In a scatter plot, the x and y numbers are used to determine where each data point will be placed. What this indicates is that the x and y coordinates of a data point intersect. Figure 11 depicts the output as a function of ultraviolet radiation. Here, we extract feature values from the feature dataset to determine the output.

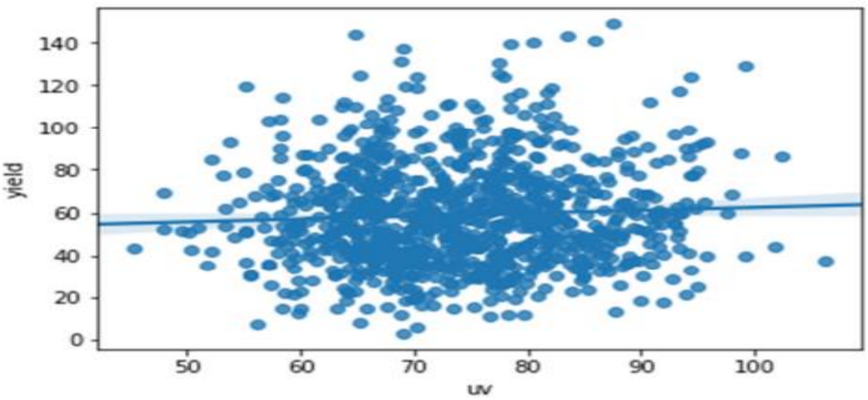

Fig 11. UV versus yield

The features' numbers are retrieved from the feature set, and the corresponding calculations are performed. Several criteria must be met before any meaningful research can be performed on the data. Parameters such as temperature, precipitation, soil type, humidity, and land size are all part of the picture. This is the data analysis and improved prediction of yield used to arrive at these characteristics. Some fields are only suitable for growing certain types of products. Some examples of these crops are wheat, rice, potatoes, sugarcane, and maize.

Table 2: Performance Results

Features	**Crops**	**Algorithm**	**Accuracy**	**Precision**
Temperature, Location, Rainfall, Soil	Wheat, Rice,	DT, ANN, RF	91%	85%
Temperature, relative humidity, Rainfall, area, pH value	Wheat, Rice, Sugarcane, corn, potato	LSTM, RNN, SVM (Proposed Algorithm)	96%	94%

In table 2, it is shown that there are specific characteristics that are used as the foundation for conducting an analysis of the dataset over a group of crops. Artificial Neural Network (ANN) and Random Forest methods are implemented, and it is found that they have 91% accuracy. In contrast, 96% accuracy is estimated using the combination of the Long-Short Term Memory (LSTM), Recurrent Neural Network (RNN), and Support Vector Machine (SVM) methods. This allows us to say categorically that the Deep Learning algorithms, in addition to the Machine Learning algorithm, play a crucial role in the process of making reliable yield predictions.

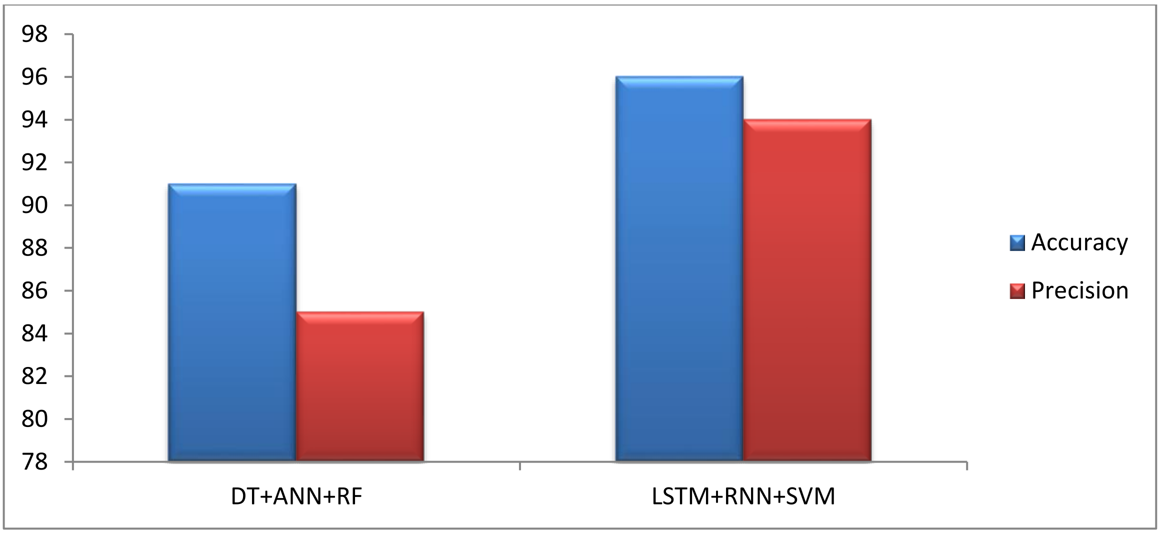

Fig12: Performance Analysis

5. Conclusion

Artificial intelligence algorithms were used in the development of the suggested model to help farmers increase crop yields and decrease financial losses caused by growing crops in unfavorable environments. The model is developed using machine learning (SVM) and deep learning (LSTM, RNN). The model determines which of several crops can be grown most efficiently on a particular plot of land by analyzing the prediction parameters. We were unable to find any other publication that used the same methods we did to predict agricultural yields. This study's findings are thus more reliable than those from previous studies that relied on alternative methods for harvest prediction. The accuracy estimates are 96%. It has a broad variety of possible future applications, is easy to implement and interface with a program that possesses a wide range of capabilities, and so on. Farmers need education, so they can use their mobile devices to access concise data on how to maximize their harvest.

There will be no future loss if the rancher is at home and the job needs to be handled at this very moment. Improvements in agribusiness will be greatly appreciated and used to aid farmers in crop output.

References

[1] M. E. Patil, M. Roshini, M. Chitrarupa, B. Laxmaiah, S. Arun and R. Thiagarajan, "A Hybrid Approach for Crop Yield Prediction using Supervised Machine Learning," 2022 8th International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 2022, pp. 1-6, doi: 10.1109/ICSSS54381.2022.9782272.

[2] A. K. Sharma and A. S. Rajawat, "Crop Yield Prediction using Hybrid Deep Learning Algorithm for Smart Agriculture," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2022, pp. 330-335, doi: 10.1109/ICAIS53314.2022.9743001.

[3] A.P.S Manideep , Dr. Seema Kharb, "A Comparative Analysis of Machine Learning Prediction Techniques for Crop Yield Prediction in India," Turkish Journal of Computer and Mathematics Education Vol.13 No.02 (2022), 120-133

[4] Maya Gopal, P.S.; Bhargavi, R, "Novel Approach For Efficient Crop Yield Prediction," Computers and electronics in agriculture, 2020

[5] Ishwarya R, Nagapooja BN, Raghavi R, Soundarya K, Prof. Chitra C, "Crop Yield Prediction Using Machine Learning Algorithm," International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:07/July-2022

[6] MA Shaik, G Manoharan, B. Prashanth, N Akhil, A Akash, *and* R S Reddy "Prediction of crop yield using machine learning," AIP Conference Proceedings **2418**, 020072 (2022)

[7] Murakami K, Shimoda S, Kominami Y, Nemoto M, Inoue S (2021) Prediction of municipality-level winter wheat yield based on meteorological data using machine learning in Hokkaido, Japan. PLOS ONE 16(10): e0258677.

[8] Rashid Ismael, H., Mohsin Abdulazeez, A., & A. Hasan, D. (2021). Comparative Study for Classification Algorithms Performance in Crop Yields Prediction Systems. *Qubahan Academic Journal*, *1*(2), 119–124.

[9] Engen, M.; Sandø, E.; Sjølander, B.L.O.; Arenberg, S.; Gupta, R.; Goodwin, M. Farm-Scale Crop Yield Prediction from Multi-Temporal Data Using Deep Hybrid Neural Networks. Agronomy 2021, 11, 2576.

[10] Klompenburg, T.V.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. Comput. Electron. Agric. 2020, 177, 105709.

[11] Shao, Y.; Ren, J.; Campbell, J.B. Multitemporal Remote Sensing Data Analysis for Agricultural Application. Compr. Remote Sens. 2018, 9, 29–38

[12] Sharma, S.; Rai, S.; Krishnan, N.C. Wheat Crop Yield Prediction Using Deep LSTM Model. arXiv 2020, arXiv:2011.01498.

[13] Cai, Y.; Guan, K.; Lobell, D.; Potgieter, A.B.; Wang, S.; Peng, J.; Xu, T.; Asseng, S.; Zhang, Y.; You, L.; et al. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. Agric. For. Meteorol. 2019, 274, 144–159.

[14] Jiang, H.; Hu, H.; Zhong, R.; Xu, J.; Xu, J.; Huang, J.; Wang, S.; Ying, Y.; Lin, T. A deep learning approach to conflating heterogeneous geospatial data for corn yield estimation: A case study of the US Corn Belt at the county level. Global change biology. Glob. Chang. Biol. 2019, 26, 1754–1766.

[15] Khaki, S.; Wang, L.; Archontoulis, S.V. A CNN–RNN Framework for Crop Yield Prediction. Front. Plant Sci. 2019, 10, 1750

[16] Khaki, S.; Pham, H.; Wang, L. Simultaneous corn and soybean yield prediction from remote sensing data using deep transfer learning. Sci. Rep. 2021, 11, 11132.

[17] Sagan, V.; Maimaitijiang, M.; Bhadra, S.; Maimaitiyiming, M.; Brown, D.R.; Sidike, P.; Fritschi, F.B. Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning. Isprs J. Photogramm. Remote Sens. 2021, 174, 265–281.

[18] Kukal, M.S.; Irmak, S. Climate-Driven Crop Yield and Yield Variability and Climate Change Impacts on the U.S. Great Plains Agricultural Production. Sci. Rep. 2018, 8, 3450.

[19] Zhang, Y.; Chipanshi, A.; Daneshfar, B.; Koiter, L.; Champagne, C.; Davidson, A.; Reichert, G.; Bédard, F. Effect of using crop specific masks on earth observation based crop yield forecasting across Canada. Remote Sens. Appl. Soc. Environ. 2019, 13, 121–137.