International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
DOI: https://doi.org/10.17762/ijritcc.v11i1.6084
***Article Received**: 28 November 2022 **Revised**: 24 December 2022 **Accepted**: 08 January 2023*

Machine Learning Based Crop Prediction on Region Wise Weather Data

Saikat Banerjee[1], Shubha Chakraborty[2], Dr. Abhoy Chand Mondal[3]
[1]Research scholar, The department of Computer science,
The University of Burdwan, West Bengal, India
saikat.banerjee56@gmail.com
[2]Research scholar, The department of Computer science,
The University of Burdwan, Golapbag, West Bengal, India
shubhac13@gmail.com
[3]HOD & Professor, The Department of Computer Science
The University of Burdwan, West Bengal, India
abhoy_mondal@yahoo.co.in

Abstract— Agriculture is a primordial occupation for human civilization, whereby farmers cultivate domesticated species of food. It refers to farming in general, which is an art and science that attempts to reform a component of the Earth's exterior through the cultivation of plants and other crops, as well as raising livestock for sustenance or other necessities for the soul and economic gain. As a result of the vital role that sustainable agriculture plays in the overall health of the nation, this sector of the economy has been the incubator for some of the most cutting-edge technological advances in recent history. Scientists and farmers have been working together to discover new methods that will allow them to increase crop production while simultaneously decreasing their water consumption and lessening their negative effects on the environment. Machine learning, deep learning, and a number of other methodologies are some examples of these approaches. A crop's expansion and maturation are both heavily influenced by the climate in which it is grown. The local climate, namely its wind speed, temperature, rainfall, and humidity, is the most exigent factor in determining the advancement or failure of crop production. If the weather is predicted prior to crop cultivation, it will be beneficial to the farmer. Machine learning is a new innovation that can solve people's real-life problems. It is a technique where a machine can act like a human and learn through experiences and the use of different types of data. Now a day, Agriculture is one of the fields of machine learning where we use different types of machine learning algorithms to predict crop production based on climate data which can benefited farmers to increase the production of the crop. In these studies, we are going to predict crop yield using LSTM based on predicted weather data.

Keywords-Machine Learning; Multiple Linear Regression; Long Short Term Memory (LSTM); Crop.

I. INTRODUCTION

The phrase "agriculture" has origins in the Latin phrase agricultra, from ager, "area," and cultra, "cultivation" or "developing." It is defined in its broadest sense, in the broadest sense of the use of natural resources "to produce life-sustaining substances, including food, fibre, forest products, horticultural plants, and their related services." Agriculture was the initial step in the evolution of mankind, and despite the enormous advancements that have taken place throughout human history, agriculture is still one of the most important human activities. Although the importance of agriculture may not be as readily apparent in certain nations as it is in others, the truth remains that every nation on the planet is dependent on agriculture in some form or another. Agriculture is the source of a significant portion of the world's raw resources, including cotton, sugar, timber, and palm oil, amongst other things [18]. These items are vital in large sectors in ways that the majority of people are completely unaware of, such as medications, diesel fuel, plastics, and a variety of other industries. Raw resources are essential to the manufacturing process; as a result, the economic well-being of a nation is heavily dependent on the quantity of raw materials available within its borders. Because of its significance to worldwide business: The majority of agriculturally produced goods that are exported to other countries are still in their early stages of development. Countries that have a plentiful supply of certain goods often export them as well as import goods that fill the gaps in their domestic supply. Prices may go up, which would wreak havoc on the country's trading patterns if the nation's agricultural sector were to falter for whatever reason [17]. When we talk about commerce, we have to mention that the majority of a developing nation's income still comes from the sale of agricultural goods. Even if industrialized nations are no longer as reliant on agriculture as they once were, the sudden cessation of all exports would nevertheless have an impact on those nations' economies. The agricultural sector continues to be one of the most important job generators, and in many places, it is even seeing expansion. Whether you see yourself as a farmer, harvester, farm mechanic, scientist, etc., there are many jobs to

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
DOI: https://doi.org/10.17762/ijritcc.v11i1.6084
***Article Received**: 28 November 2022 **Revised**: 24 December 2022 **Accepted**: 08 January 2023*

be had in this subject [19]. In developing countries, agricultural activities are helping to reduce the high levels of unemployment. Concerning the poverty discount, the evidence shows that focusing on agriculture is more effective than investing in different regions [20].

The advent of artificial intelligence (AI) has brought about a shift in how we conceptualize contemporary science. John McCarthy, a well-known American computer scientist, came up with the concept that would later be known as "artificial intelligence" in the summer of 1956 at the Dartmouth conference. AI has been taken to another level by the profound British mathematician Alan Turing. He introduced a machine called an imitation game, also widely known as a Turing machine, in the year 1950. If a machine can pass the Turing test, then the machine is called an intelligent machine. Another famous test is the Chinese Room, which was proposed by American philosopher John Rogers Searle. [9] Artificial intelligence covers an extensive range in the branch of computer science where a computer or computer-controlled robot performs a task like a human. Intelligent entities or agents need to be able to do both mundane tasks such as planning routes, recognising people or objects through vision, communicating through natural language, etc., and expert tasks such as medical diagnosis, mathematical problem solving, etc. If a machine is intelligent, then it has some intelligent behaviors like perception, reasoning, learning, understanding languages, and solving problems. There are four types of approaches to AI: strong AI, weak AI, applied AI, and cognitive AI. Strong AI aims to build machines that have self-aware consciousness and can solve problems. Here, a machine would have the same intelligence as a human [8]. Weak AI is also known as narrow AI because it is focused on a single task. This type of AI is used in our daily lives, such as in Apple's Siri, Amazon's Alexa, Google Assistant, etc. Applied AI deals with a modern smart system such as a face-recognized security system. In cognitive AI, computers are used to test theorems about how the human mind works when completing a particular task. In a survey of 2017, among the five companies, one used artificial intelligence. In the year between 2015 and 2019, the total amount of research into AI increased by 50%[9]. There are many applications of AI, such as speech recognition, customer service, computer vision, recommendation engines, automated stock trading, etc. Some of the fields where AI is used are healthcare, business, agriculture, education, finance, law, manufacturing, banking, transportation, security, astronomy, gaming, robotics, entertainment, etc. At present, one of the famous parts of artificial intelligence is machine learning, which is widely used all over the world [11].

In the field of artificial intelligence known as machine learning, statistical methods are used to teach machines how to get better with experience. This allows computers or computer-controlled robots to act and make decisions based on data, as opposed to being definitively programmed to carry out a specific task. In the year 1959, the well-known American IBMer and pioneer (in the fields of computer games and artificial intelligence) Arthur Samuel came up with the phrase "machine learning (ML)." Machine learning is utilized in the contemporary world for primarily two purposes: the first is to categorise the data based on certain machine learning models that have been constructed, and the second is to predict future outcomes based on these models. The process of feature extraction in machine learning is performed by hand. In machine learning, there are three different kinds of approaches: supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, it uses labelled datasets to train the algorithms that help classify data or predict outcomes accurately. [13] The model is able to monitor its own precision and improve over time because of the marked inputs and outputs it receives. In simple words, we first train the machine with the training input and corresponding desired output, and then we predict the output using the test datasets. There are two distinct categories of issues that might arise with supervised learning: classification and regression. To solve the classification problem, a classification algorithm is used where the output variable is categorical. Some real-world examples are spam detection, speech recognition, handwriting recognition, document classification, biometric identification, etc. Some popular classification algorithms are Random Forest, Decision Tree, Logistic Regression, Support Vector Machine, etc. [14]. Regression allows us to predict by learning the relationship between the features of our data and some of the continuous-valued responses. This method is used to predict continuous variables, such as market trends, weather predictions, etc. Some famous algorithms that are used in regression are simple linear regression, multivariate regression, decision trees, lassa regression, etc. Inside the procedure of unsupervised learning, only the input data are considered, with no associated output variable [5]. The main goal here is to create a model that accurately represents the underlying structure or distribution of the data in order to learn more about it. That means the algorithm is to group or categories the unlabeled datasets according to the patterns, similarities, or differences [6]. Unsupervised learning is classified into two types: clustering and association. Clustering is a technique where similar data is grouped into a cluster, whereas different data sets are grouped in another cluster. Companies use this algorithm to group customers based on their purchasing habits. Some clustering algorithms are k-means clustering, mean-shift algorithm, DBSCAN, principal component analysis, independent component analysis, etc. Association is a technique that finds the relationships among a large dataset. The main goal of this algorithm is to find the correlation between data and map those data in order to maximize profit [12]. These algorithms are mainly used in

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
DOI: https://doi.org/10.17762/ijritcc.v11i1.6084
***Article Received**: 28 November 2022 **Revised**: 24 December 2022 **Accepted**: 08 January 2023*

market basket analysis, web usage mining, continuous production, etc. Some of the association learning algorithms are the apriori algorithm, Eclat, FP-growth algorithm, etc. In reinforcement learning, it works on a feedback-based system. It allows software agents to automatically determine the ideal behavior within a specific context to maximize their performance. Reinforcement learning is about the reaction between two elements, the first is the environment, and the second is the agent of learning [9]. Here, the agent gets rewarded for every good action, whereas every bad action gets punished. Agents receive the award and punishment as feedback to improve themselves. The reinforcement learning method is similar to how humans learn; for example, a child learns different types of things by experiencing them in day-to-day life. There are two types of reinforcement learning: positive reinforcement learning and negative reinforcement learning. Positive reinforcement learning increases the strength and frequency of the behavior when an event occurs due to that behavior. Negative reinforcement learning increases the strength of behaviors because negative conditions are avoided [10]. Some of the applications of reinforcement learning are video games, resource management, robotics, text mining, etc. Deep learning symbolizes a shift in the conventional wisdom of machine learning [12].

The subfield of machine learning known as deep learning was developed in which feature extraction is carried out using a neural network. On the other hand, deep study is simply a system for acquiring knowledge that is stimulated by the shape of the human mind. Through constantly analyzing data with a predetermined logical framework, "deep learning" computers make an effort to arrive at conclusions that are comparable to those that people would reach [2]. Deep learning accomplishes this goal via the use of neural networks, which are multi-layered structures made up of algorithmic components. The structure of the human mind served as inspiration for the organization of the neural network's architecture [8]. Simply because our minds are used to recognizing patterns and organize things. Neural networks can be taught to perform the same tasks on different types of data. The character layers of neural networks may also be thought of as a form of filtering that moves from large to diffused, improving the likelihood of detecting and producing an appropriate result. The human brain works similarly [11]. Every time we acquire new statistics, the brain tries to evaluate them with recognized objects. Deep neural networks, which are similar to traditional neural networks, also employ the same general concept. The use of neural networks enables us to carry out a variety of tasks, such as grouping, categorizing, and performing regression analysis [4]. Through the use of neural networks, we are able to classify or categories unlabeled records in a manner that is congruent with the similarities that exist between the examples in this data. Alternatively, within the confines of the category case, we can educate the community on a classified dataset and classify the samples on this dataset into distinct classes. Deep learning models can tackle problems that device mastery models cannot. Deep learning has led to all recent AI developments [13, 16]. Without deep learning, technologies such as self-driving vehicles, chatbots, and artificially intelligent companions like Alexa and Siri just wouldn't exist. Netflix and Youtube may not know which movies and TV shows we prefer, and Google Translate may remain as rudimentary as it was 10 years ago (before Google turned to neural networks for this app). These technologies use neural networks. Even now, artificial neural networks and deep learning are driving a new industrial revolution. Deep learning is our best and most evident way to actually increase system intelligence [21]. Statistics power it, which explains its popularity. Deep learning advances will abound in the "big data generation" generation. Deep Learning models improve as more training data is added, whereas SVM and Naive Bayes classifiers reach a saturation point [14].

II. Related Work

1. In this paper, the author introduces a crop yield prediction model that helps farmers produce crops using area, yield, production, area under irrigation, and rainfall. Here the author used four types of machine learning techniques such as decision trees, linear regression, lasso regression, and ridge regression. The datasets taken for this experiment were collected from government and private websites. Following the application of these four techniques, decision trees provide the highest accuracy of 98.62 percent.
2. The author discusses a method that utilises supervised machine learning for predicting crop yields in the agricultural industry. All of these factors ph, humidity, temperature, rainfall, humidity, and crop name—are forecasted on the basis of previous data. The dataset that the author used was collected from Kaggle. The random forest machine learning technique was used here. After using this technique, the author achieves an acceptable result with good accuracy.
3. In this study, the author introduced a deep learning-based weather forecast system. For this system, the Python Keras library and the Pandas library are used. The dataset is collected by the National Oceanic and Atmospheric Administration (NOAA) and the National Climatic Data Center (NCDC). The authors construct the neural network using the Keras Sequential Model as their deep learning model. After using this model, the result is almost accurate.
4. The author proposed a crop yield prediction using a machine learning algorithm based on climate, which

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
DOI: https://doi.org/10.17762/ijritcc.v11i1.6084
Article Received: 28 November 2022 Revised: 24 December 2022 Accepted: 08 January 2023

helps farmers produce crops. Logistic regression, Naïve Bayes and Random Forest techniques are used here. The data are gathered from different sources, such as data.gov.in, indiawaterportal.org, and power.larc.nasa.in. Libraries are Scikit-Learn, Numpy, Keras, and Pandas. Random Forest provides the highest level of accuracy.

5. In this paper, the author describes a LSTM -based Recurrent Neural Network(RNN) based crop prediction model using deep learning based on weather. The author constructed the model based on the performance data obtained from the Uniform Soybean Tests (UST) conducted in North America. This model gives a good result, which is helpful to predict the crop.
6. The author introduced a deep neural network based short-term local weather forecast using dense weather stations. For precipitation forecasts, radar echo is often used as the source for the prediction of the author presented a short-term local weather prediction that was based on a deep neural network and made use of a dense network of weather stations. In the process of making predictions about precipitation, radar echo is frequently employed as the primary source in a LSTM-based precipitation nowcast model. The author of this study presents deep learning approaches as a means of predicting weather aspects using extensively observed meteorological data gathered from POTEKA. According to the results of our research, deep neural networks provide the best level of accuracy for rain forecasting when compared to other machine learning approaches. In the instance of the "No-Rain" forecast, the results produced by the block-type network are more accurate than those produced by the fully linked network.
7. In this paper, the author proposes a crop prediction model using deep neural networks. The dataset included 2267 experimental hybrids planted in 2247 locations between 2008 and 2016 across the United States and Canada. For this deep neural network, the author used the Tensorflow open source software library in Python. This model gives a good result with good accuracy.
8. The author presents a strategy for predicting agricultural output that utilises both deep and machine learning as its two primary components. SVM is employed in this situation as an example of a machine learning algorithm, whereas LSTM and RNN are examples of deep learning algorithms. The dataset for this model was obtained from Kaggle. After implementing this model, the author observed that the model gave a good result with 97% accuracy.

III. PROPOSED MODEL

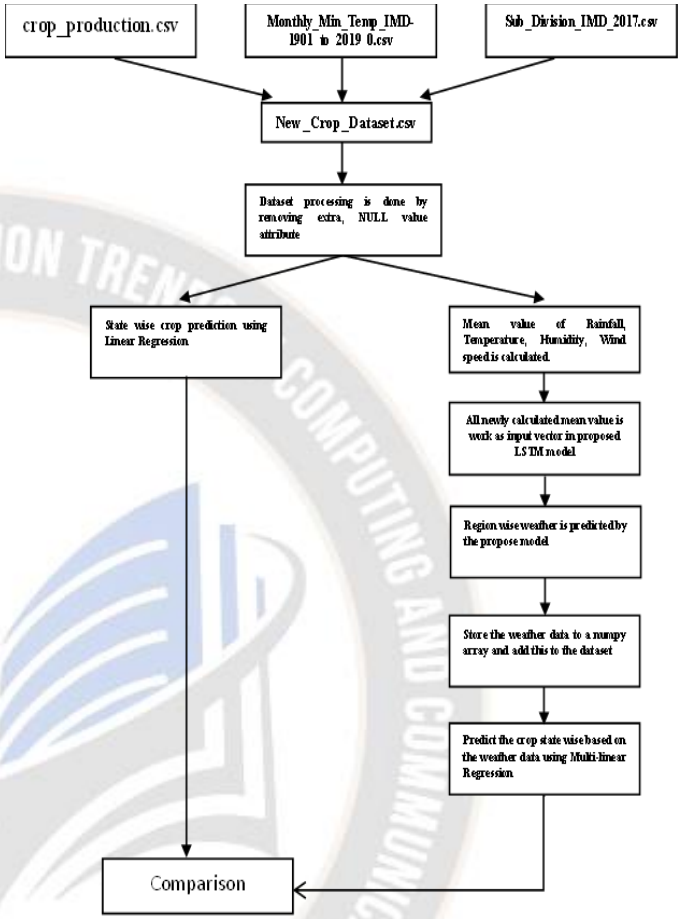

Figure1: Step by Step Workflow of Our Proposed Work

A. Methodology

To naturalize the precepting of the framework in this research, we briefly preface the constituency of multiple linear regression and RNN-LSTM and, in addition, recapitulate why LSTM should be promoted in forecast analysis. Even if it has been shown that RNNs are excellent for sequential prediction problems, it could still be challenging to recognize the long-term reliance of the data. This is essentially the outcome of the explosion/vanishing gradient issue, which arises when the gradients propagating in the RNN are applied to additional layers. As more layers are applied, the problem becomes more severe. Datasets used

We have used and located three datasets from the official website of the Indian government, and we have used these datasets in our research. The datasets came from many different public and commercial sources, as well as official government ones.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
DOI: https://doi.org/10.17762/ijritcc.v11i1.6084
***Article Received**: 28 November 2022 **Revised**: 24 December 2022 **Accepted**: 08 January 2023*

- Crop_production.csv is a Kaggle dataset that contains 12 years of data categorized by crop, crop year, production, area, state, district, and season. [15]
- Temperature: From data.gov.in, 12 years of temperature data are organized by month and district in India.[16]
- Rainfall: 12 years of rainfall data from data.gov.in are organized by month and district in India.[16]

The combined dataset contains 142 000 fields. It covers 23 states throughout India and provides information such as area, temperature, crop name, rainfall, humidity, wind speed, and productivity for each state. The production, temperature, rainfall, humidity, and wind speed datasets were denominated on the basis of seasons. This data was then associated according to the states in the final dataset. The data analysis stage resulted in an accurate dataset. Table 1 depicts a portion of the final dataset.

State_Name	District_Name	Crop_Year	Season	Crop	Area	Production	ANNUAL-Rainfall	ANNUAL-Temperature	Humidity(%)	W
Andhra Pradesh	ANANTAPUR	2004	Kharif	Arhar/Tur	32752	7893	2890	19.51	77	
Assam	SONITPUR	2014	Annual	Potato	8858	62136	2711.7	19.44	55	
Bihar	AURANGABAD	2014	Rabi	Masoor	13612	9474	1158	19.25	71	
Chhattisgarh	JANJGIR-CHAMPA	2010	Kharif	Maize	396	687	872.7	19.22	55	
Gujarat	BANAS KANTHA	2014	Kharif	Bajra	173000	108300	622.9	19.03	87.5	
Himachal Pradesh	KULLU	2016	Rabi	Garlic	569	458	3141.1	19.51	44.9	
Kamataka	BAGALKOT	2014	Rabi	Sunflower	62717	35641	3524	19.08	57.8	
Kamataka	BAGALKOT	2016	Rabi	Wheat	21202	32903	2757.5	19.09	60.2	
Kamataka	BAGALKOT	2015	Summer	Maize	1509	5691	2082	19.13	88	
Kamataka	BAGALKOT	2017	Summer	Sunflower	4087	3052	3663.9	19.01	85.6	
Madhya Pradesh	HOSHANGABAD	2013	Annual	Sesamum	57	18	838	19.31	68	
Madhya Pradesh	HOSHANGABAD	2018	Annual	Sugarcane	984	31620	644.5	19.27	71.65	
Odisha	JAGATSINGHAPUR	2016	Winter	Rice	92000	168000	1520	19.09	58	
Odisha	JAGATSINGHAPUR	2015	Winter	Sugarcane	489	33769.9	919.1	19.41	63	
Rajasthan	JALORE	2017	Annual	Onion	253	698	355.7	19.64	69.12	
Rajasthan	JALORE	2017	Annual	Potato	77	146	251.5	19.34	72.32	
Uttar Pradesh	KAUSHAMBI	2018	Rabi	Maize	11	51	185.9	19.02	53	
Uttar Pradesh	KAUSHAMBI	2014	Rabi	Potato	3363	36283	454.8	19.02	59	
West Bengal	HOOGHLY	2014	Annual	Potato	87976	1807873	661.7	19.37	73.4	
West Bengal	BARDHAMAN	2015	Winter	Rice	417180	1365492	503.7	19.07	53.1	
West Bengal	BARDHAMAN	2016	Autumn	Rice	13530	43731	571.3	19.54	61.2	

Figure2: New Crop Dataset

IV. MULTIPLE LINEAR REGRESSION

Multiple regression, often shortened to "many regressions," is a mathematical process for predicting the values of a response parameter by considering a number of potential causes. Alternative names for this technique include multiple linear regression. The purpose of the statistical method is to graphically display the linear connection between the explanatory (independent) and dependent (dependent) variables (the established variables). More than one extrapolation may be seen as an expansion of the partial least squares (OLS) approach, since it considers a larger number of factors to explain the data.

$$yp_i = a_0 + a_1 x_{i1} + a_2 x_{i2} + \cdots + a_n x_{in} + \epsilon \qquad (1)$$

where,
yp_i=dependent variable
x_i=explanatory variables
a_0=y-intercept (constant term)
a_n=slope coefficients for each explanatory variable
ϵ= the term for the model's inaccuracy (also called as the residuals).
The following figure shows state wise crop production.

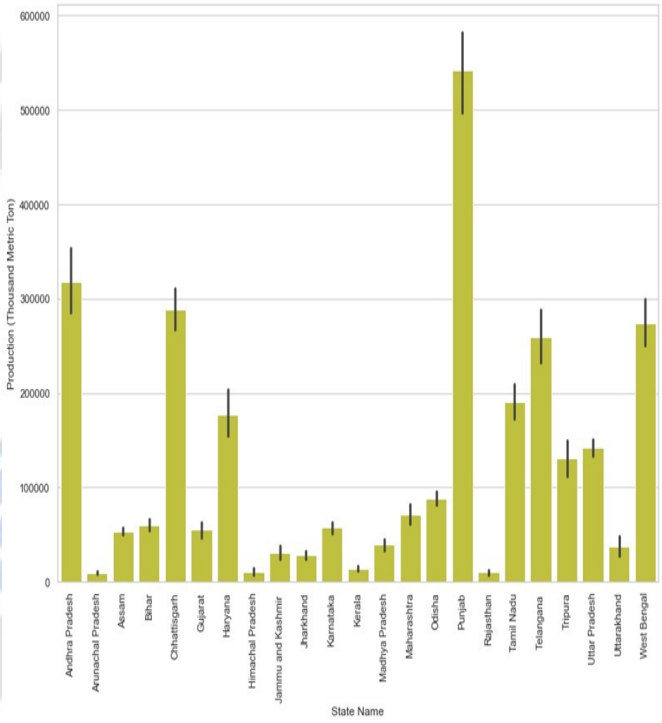

Figure3: State wise Crop Production

V. RNN

Recurrent neural networks, also known as RNNs, are a category of neural networks that permit previous outputs to be used as inputs at the same time as having hidden states. They are typically as follows: For each time step n, the activation cn and the output on are expressed as follows:

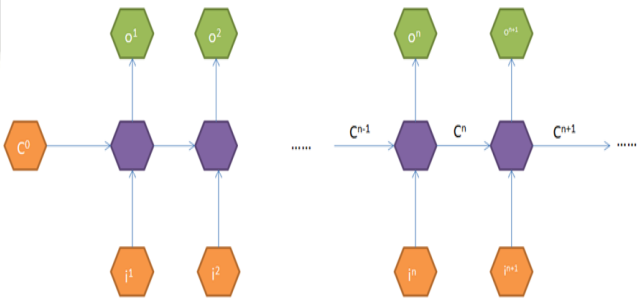

Figure4: RNN Model

$$c^n = f_1(w_{cc}c^{n-1} + w_{ci}i^n + b_c) \qquad (2)$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
DOI: https://doi.org/10.17762/ijritcc.v11i1.6084
Article Received: 28 November 2022 Revised: 24 December 2022 Accepted: 08 January 2023

$$o^n = f_2(w_{oc}c^n + b_o) \qquad (3)$$

where w_{cc}, w_{ci}, w_{oc}, b_c and b_o are coefficients that are shared temporally, and f_1, f_2 are the activation functions. RNNs produce either a graph with no order or an ordered graph with a time sequence. This shows temporal dynamic behaviour. RNNs, which are a kind of feed-forward neural network, have the ability to interpret keystreams of different lengths by using their memory. They are useful for segmentation, linked handwriting popularity, and speech popularity. RNNs are theoretically Turing is full and can process any input sequence with any package. "Recurrent neural communities" refer to infinite impulse response networks, whereas "convolutional neural communities" relate to finite impulse response networks. Network instructions are temporary. Finite impulse recurrent networks are directed acyclic graphs that can be unrolled and altered using a purely feed-forward neural network, whereas unlimited ones are directed cyclic graphs that cannot. The neural network may directly regulate storage in finite impulse and endless stimulus recurrent networks. Any network or graph with time delays or feedback loops may replace storage. Gated states, also known as gated memories, are part of long-term memory networks and gated recurrent devices. Neural community.

Backpropagation via Time: In system mastery, backpropagation is known as a workhorse collection of rules. Calculating the gradient of an error feature with recognition of a neural network's weights may be accomplished via the process of back propagation. In order to locate the partial derivative of the mistakes with regard to the weights, the set of rules works its way in the opposite direction through the several levels of gradients. After then, the back prop will employ those weights to reduce the margins of error when training. At each and every moment in time, back propagation is carried out. The expression for the loss's derivative with regard to the weight matrix w at time step N is as follows:

$$\frac{\Delta\alpha^N}{\Delta w} = \sum_{n=1}^{N} \frac{\Delta\alpha^N}{\Delta w} \qquad (4)$$

The disappearing gradient phenomenon and the expanding gradient phenomenon are both common occurrences within the framework of RNNs. It is much more difficult to capture long-term relationships due to multiplicative gradients, which may be exponentially decreasing or expanding with regard to the number of layers. This is the root cause of their occurrence. Kinds of gates: so as to remedy the vanishing gradient hassle, particular gates are used in a few styles of RNNs and usually have a well-described cause.

$$G = \delta(Wa^n + Ux^{n-1} + b) \qquad (5)$$

where W, U, and b are coefficients specific to the gate, and δ is the sigmoid function. Update gate When it arrives to GRU and LSTM, the value of Gu is What factors affect the current relevance of numerous events that occurred in the past. Relevance gate Gr is also used in both GRU and LSTM to drop previous information. Forget gate Gf is used in LSTM to erase a cell or not. Output gate Go is also used in LSTM to determine how much to reveal about a cell.

A. LSTM

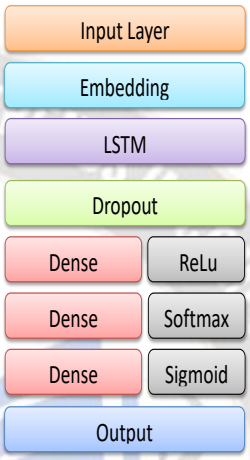

Figure5: LSTM Workflow

Long short-term memory, also known as LSTM, is a structure of artificial recurrent neural networks (RNNs) that is used within the context of the study of deep learning. LSTM was first proposed by Sepp Hochreiter and Jürgen Schmidhuber. Unlike trendy feed forward neural networks, LSTM has comment connections. It is able to handle not only individual data sources but also whole record sequences. LSTM is useful for unsegmenting, linked handwriting reputation, voice popularity, and network visitor or IDS anomaly detection (intrusion detection structures). Cells, input gates, output gates, and neglect gates make up LSTM units. Because of the three gates, the flow of information to and from the cell, which stores values for arbitrary periods of time, may be altered. Time series records may have unpredictable latency between important occurrences, making LSTM networks ideal for identifying, analyzing, and predicting them. LSTMs can now handle the "vanishing gradient" issue while training classical RNNs. LSTM has a length-insensitive advantage over RNNs, hidden Markov models, and other series learning algorithms in severe situations.

The dropout layer helps avoid overfitting by randomly setting input device charges to zero at some point during the training process. This is done with a frequency of charge at each step. Inputs that are not set to 0 are scaled up using 1/(1 - charge) to keep the total value constant across all inputs. Dropout is a type of regularization in neural networks that allows for a reduction in the amount of interdependent learning that occurs

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
DOI: https://doi.org/10.17762/ijritcc.v11i1.6084
***Article Received**: 28 November 2022 **Revised**: 24 December 2022 **Accepted**: 08 January 2023*

between the neurons. If the layer that comes before it has a lot of connections to it, then the layer that comes after it is considered to be dense. This suggests that the neurons that make up the thick layer are linked to each and every neuron that was present in the layer that came before it. This residue is the layer that is used the most often in synthetic neural network networks' highest layer, which is the lattice of positive.

In one implementation, the neurons of the dense layer are accountable for performing matrix-vector multiplication. In this version, the dense layer neuron gets output from each neuron in its preceding layer. Through the process of matrix vector multiplication, It is possible to ensure that the row matrices of the result from the previous stages have the same value as the matrix form of both the dense surface. This can be done by ensuring that the product from the stages that arrived before it has a dense layer in between them. For combining matrices and vectors, the fundamental rule that should be adhered to is that each row matrix must consist of the same number of columns as the matrix form. This is the case because combining matrices and vectors might provide unexpected results.

VI. RESULTS AND ANALYSIS

In this section, our research produces a set of results accepted from experiments directed to compare the performance of twelve-year data using a proposed RNN-LSTM and multilevel regression framework for performing two major tasks: region-wise weather prediction and based on predicted weather data crop analysis.

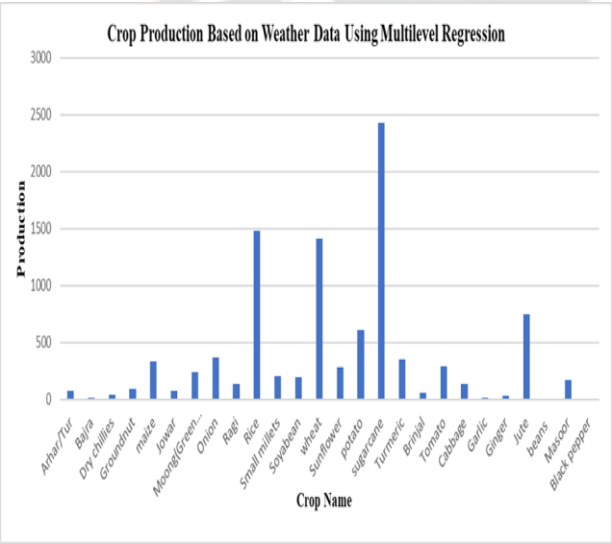

Figure6: Crop Production Prediction

The results of a multilevel regression analysis are shown in figure 6, which can be seen here. We begin the study by processing the climatic information, which serves as supervised data for the development of the model. This is the first step in the research. The results from the initial stage are obtained by applying the conceptual approach to the experiment that was carried out in the first phase of the research project.

The results of a multilevel regression analysis are shown in figure 6, which can be seen here. We begin the study by processing the climatic information, which serves as supervised data for the development of the model. This is the first step in the research. The results from the initial stage are obtained by applying the conceptual approach to the experiment that was carried out in the first phase of the research project. As seen in Figure 7, the crop production based on weather data.

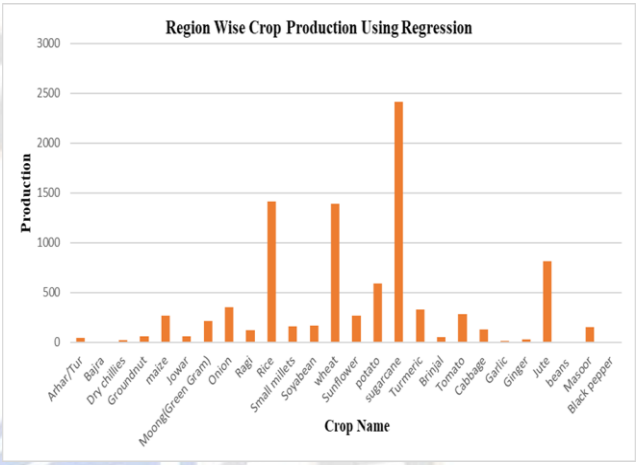

Figure7: Crop Production Prediction based on Weather Data

VII. CONCLUSION

Table 1: Comparison Table

Crop Name	Algorithm Used	Accuracy
Arhar/Tur	Multiple linear regression	88.21
	LSTM	90.56
Bajra	Multiple linear regression	84.66
	LSTM	88.76
Castor seed	Multiple linear regression	93.33
	LSTM	97.42
Cotton(lint)	Multiple linear regression	90.25
	LSTM	94.75
Dry chillies	Multiple linear regression	86.75
	LSTM	91.12
Groundnut	Multiple linear regression	87.31
	LSTM	92.04
Horse-gram	Multiple linear regression	88.98
	LSTM	93.24
Jowar	Multiple linear regression	87.35
	LSTM	91.88
Korra	Multiple linear regression	88.08

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
DOI: https://doi.org/10.17762/ijritcc.v11i1.6084
***Article Received**: 28 November 2022 **Revised**: 24 December 2022 **Accepted**: 08 January 2023*

	LSTM	92.55
Maize	Multiple linear regression	88.31
	LSTM	93.07
Moong(Green Gram)	Multiple linear regression	89.27
	LSTM	93.62
Other Kharif pulses	Multiple linear regression	89.65
	LSTM	93.92
Ragi	Multiple linear regression	87.47
	LSTM	92.12
Rice	Multiple linear regression	90.22
	LSTM	95.14
Sugarcane	Multiple linear regression	90.32
	LSTM	95.04
Sunflower	Multiple linear regression	90.25
	LSTM	94.88

At first, we predicted crop production state wise from this dataset using linear regression. Following that, we used LSTM to forecast weather for each state. With this result, we have predicted crop production using multilinear regression. For each crop, the accuracy of the algorithm's prediction performance is displayed in Table I, which demonstrates how effective the algorithm is as a prediction mode The results of a multilevel regression analysis are shown in figure 6, which can be seen here. We begin the study by processing the climatic information, which serves as supervised data for the development of the model. This is the first step in the research. The results from the initial stage are obtained by applying the conceptual approach to the experiment that was carried out in the first phase of the research project. As seen in Figure 7, . Due to collecting data from various datasets, there is uncertainty in our result. If we had an IoT device, our result would have been much better.

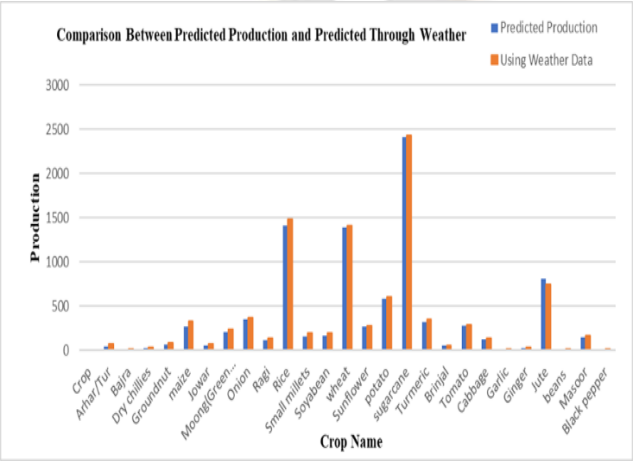

Figure8: Compare Between Crop Production and Crop Production Based on Weather Data

It outperforms both on average, as shown in Figure 8's accuracy comparison to crop production and crop production based on weather data. When comparing the suggested model to the existing one, it is necessary to establish whether there is a significant increase in terms of accuracy. After comparing two results, we have reached the conclusion that the prediction of crop production using weather data is 92% better. This study proposes a system that will provide farmers with production estimates based on meteorological conditions and area under cultivation. Using this, a farmer can decide whether to produce a specific crop or choose an alternative crop in the event that yield forecasts are negative.

REFERENCES

[1] Ms Kavita, Pratistha Mathur, "Crop Yield Estimation in India Using Machine Learning", 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA) Galgotias University, Greater Noida, UP, India. Oct 30-31, 2020.

[2] Dr. Y. Jeevan Nagendra Kumar, V. Spandana, V.S. Vaishnavi, K. Neha, V.G.R.R. Devi, "Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector", Proceedings of the Fifth International Conference on Communication and Electronics Systems (ICCES 2020) IEEE Conference Record # 48766; IEEE Xplore ISBN: 978-1-7281-5371-1.

[3] Jarrett Booz, Wei Yu, Guobin Xu, David Griffith, Nada Golmie, "A Deep Learning-Based Weather Forecast System for Data Volume and Recency Analysis" , 2019 International Conference on Computing, Networking and Communications (ICNC): Green Computing, Networking, and Communications.

[4] Anakha Venugopal, Aparna S, Jinsu Mani, Rima Mathew, Prof. Vinu Williams, "Crop Yield Prediction using Machine Learning Algorithms", International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Published by, www.ijert.org NCREIS - 2021 Conference Proceedings.

[5] Shook J, Gangopadhyay T, Wu L,Ganapathysubramanian B, Sarkar S, Singh AK, "Crop yield prediction integrating genotype and weather variables using deep learning" , PLoSONE 16(6): e0252402. https://doi.org/10.1371/journal.pone.0252402, June 17, 2021.

[6] Kazuo Yonekura, Hitoshi Hattori and Taiji Suzuki, "Short-term local weather forecast using dense weather station by deep neural network", 2018 IEEE International Conference on Big Data (Big Data).

[7] Khaki S and Wang L, "Crop Yield Prediction Using Deep Neural Networks", Front. Plant Sci. 10:621.doi: 10.3389/fpls.2019.00621, 22 May 2019.

[8] Sonal Agarwal, Sandhya Tarar, " A Hybrid Approach for Crop Yield Prediction Using Machine Learning and Deep Learning Algorithms", CONSILIO 2020 Journal of Physics: Conference Series , 1714 (2021) 012012 IOP Publishing doi:10.1088/1742-6596/1714/1/012012.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
DOI: https://doi.org/10.17762/ijritcc.v11i1.6084
***Article Received**: 28 November 2022 **Revised**: 24 December 2022 **Accepted**: 08 January 2023*

[9] Rishi Gupta, Akhilesh Kumar Sharma, Oorja Garg, Krishna Modi , Shahreen Kasim, Zirawani Baharum, Hairulnizam Mahdin, Salama A. Mostafa, "WB-CPI: Weather Based Crop Prediction in India Using Big Data Analytics ", VOLUME 9, IEEE Access, 2021.

[10] Thayakaran Selvanayagam, Suganya S, Puvipavan Palendrarajah Mithun Paresith Manogarathash, Anjalie Gamage, Dharshana Kasthurirathna , "Agro-Genius: Crop Prediction using Machine Learning" , Volume 4, Issue 10, International Journal of Innovative Science and Research Technology, October 2019.

[11] Sonal Jain, Dharavath Ramesh, "Machine Learning convergence for weather based crop selection" , IEEE International Students' Conference on Electrical, Electronics and Computer Science, 2020.

[12] Shubham Madan1, Praveen Kumar, Seema Rawat, Tanupriya Choudhury, "Analysis of Weather Prediction using Machine Learning & Big Data" , 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE-2018) Paris, France 22-23 June 2018.

[13] Urmay Shah, Sanjay Garg, NehaSisodiya, Nitant Dube, Shashikant Sharma, "Rainfall Prediction: Accuracy Enhancement Using Machine Learning and Forecasting Techniques", 5th IEEE International Conference on Parallel, Distributed and Grid Computing(PDGC-2018), 20-22 Dec, 2018, Solan, India.

[14] Aruvansh Nigam, Saksham Garg, Archit Agrawal, Parul Agrawal, "Crop Yield Prediction Using Machine Learning Algorithms" , 2019 Fifth International Conference on Image Information Processing (ICIIP).

[15] Online available: https://www.kaggle.com/datasets/

[16] Online Available: https://pib.gov.in/

[17] Online Available: https://www.kisaanhelpline.com

[18] Online Available: https://agricoop.nic.in/#gsc.tab=0

[19] Online Available: https://iasri.icar.gov.in/

[20] Online Available: https://krishi.icar.gov.in/

[21] Jitendra Kumara, Rimsha Goomer, Ashutosh Kumar Singh, "Long Short Term Memory Recurrent Neural Network (LSTM-RNN) Based Workload Forecasting Model For Cloud Datacenters", Long Short Term Memory Recurrent Neural Network (LSTM-RNN) Based Workload Forecasting, 2017.