International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Agriculture Recommendation Using ML and DL techniques

Mrs.B. Aruna Sri[1], M.A.Haritha[2], M.A.Harika[3]
[1]Assistant Professor, Department of CSE, Malla Reddy Engineering College for Women, Hyderabad, Telangana, India.
[2,3]UG-Students, Department of CSE, Malla Reddy Engineering College for Women, Hyderabad, Telangana, India.

ABSTRACT

Agriculture is the pillar of the Indian economy and more than 50% of India's population are dependent on agriculture for their survival. Variations in weather, climate, and other such environmental conditions have become a major risk for the healthy existence of agriculture. Machine learning (ML) plays a significant role as it has decision support tool for Crop Yield Prediction (CYP) including supporting decisions on what crops to grow and what to do during the growing season of the crops. The present research deals with a systematic review that extracts and synthesize the features used for CYP and furthermore, there are a variety of methods that were developed to analyze crop yield prediction using artificial intelligence techniques. The major limitations of the Neural Network are reduction in the relative error and decreased prediction efficiency of Crop Yield. Similarly, supervised learning techniques were incapable to capture the nonlinear bond between input and output variables faced a problem during the selection of fruits grading or sorting. Many studies were recommended for agriculture development and the goal was to create an accurate and efficient model for crop classification such as crop yield estimation based on the weather, crop disease, classification of crops based on the growing phase etc., This paper explores various ML techniques utilized in the field of crop yield estimation and provided a detailed analysis in terms of accuracy using the techniques.

INTRODUCTION

Agriculture is the backbone of India's economy since its plays a vital role in the survival of every human and animal in India [1]. The worldwide population was estimated at 1.8 billion in 2009 and is predicted to increase to 4.9 billion by 2030, leading to an extreme increase in demand for agricultural products. In the future, agricultural products will have higher demand among the human population, which will require efficient development of farmlands and growth in the yield of crops. Meanwhile, due to global warming, the crops were frequently spoiled by harmful climatic situations [2]. A single crop failure due to lack of soil fertility, climatic variation, floods, lack of soil fertility, lack of groundwater and other such factors destroy the crops which in turn affects the farmers. In other nations, the society advises farmers to increase the production of specific crops according to the locality of the area and environmental factors [3]. The population has been increasing at a significantly higher rate, so the estimation and monitoring of crop production is necessary [4]. Accordingly, an appropriate method needs to be designed by considering the affecting features for the better selection of crops with respect to seasonal variation [5].

The core objective of crop yield estimation is to achieve higher

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

agricultural crop production and many established models are exploited to increase the yield of crop production. Nowadays, ML is being used worldwide due to its efficiency in various sectors such as forecasting, fault detection, pattern recognition, etc. The ML algorithms also help to improve the crop yield production rate when there is a loss in unfavorable conditions. The ML algorithms are applied for the crop selection method to reduce the losses crop yield production irrespective of distracting environment.

so the estimation and monitoring of crop production is necessary [4]. Accordingly, an appropriate method needs to be designed by considering the affecting features for the better selection of crops with respect to seasonal variation [5].

1. PURPOSE OF THE PROJECT

1.Depending on the dissimilar crop feature divisions, the modulating factor values of ML algorithms differ to attain perfect approximation.

2.When the quantity of input elements is reduced, ANN is utilized. The optimal feature was being empirically selected for appropriate crop yield estimation.

3.The advantage of ML method regression is to avoid difficulties of using a linear function in large output sample space and optimization of complex problems transformed into simple linear function optimization.

4.ML algorithm can be executed with an enormous soil dataset for crop yield estimation.

1. EXISTING SYSTEM

1. Tseng [2] utilized intelligent agriculture Internet of Things (IoT) equipment to monitor the crop yield prediction. The crops were generally damaged by weather conditions and the existing models used big data in intelligent agriculture to predict the crop yield farm. The developed model utilized an IoT sensor device that monitored the overall agricultural farm and sensed the atmospheric pressure, humidity, moisture content, temperature and soil salinity. The objective of big data analysis in IoT was to analyze and understand crop growing methods practiced by the farmers along with examining environmental deviations. An advantage of the developed model was 3D cluster evaluated the relation between environmental factors and subsequently examined the guidelines obtained from the farmers. However, the developed model showed unusual distribution when it was exposed to potential risk in air humidity, soil moisture content, and temperature.

Tiwari and Shukla [7] developed a model for crop yield Prediction by using CNN and Geographical Index. The existing model faced a problem during a continuous breakdown in agricultural drifts for crop cultivation which were not suitable with environmental factors like temperature, weather and soil condition. The developed CNN model which used spatial features as input were trained by BPNN for error prediction. An advantage of the developed model was that it was implemented on a real-time dataset that was taken from authentic geospatial resources. However, the developed model reduced the

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

relative error but decreased the efficiency of crop yield prediction

DISADVANTAGES

- No efficient methods are used.
- No real time data is used.
- More complex.

PROPOSED SYSTEM

The developed model introduces a deep meta- architecture to predict the pests in plants. The developed model considers three key features of indicators: Single Shot Multibox Detector (SDD), Faster region-based CNN and Region-Based Fully CNN, which is known as deep meta-architecture. The execution of the deep meta-architecture and feature extractors furthermore suggested a method for a global and local period explanation. The data growth increases the precision and also reduced the number of false positives in training. The benefit of the developed model was it successfully identified different kinds of pests and diseases by dealing with complex situations from a nearby area. Due to the usage of complex pre-processing techniques, the robust deep learning method consumes more time and high computational price.

Sun et al. [12] utilized the Deep CNN-LSTM method to predict the soybean yield estimation. The Yield prediction was an immense consequence for yield mapping, harvest management, crop insurance, crop market planning, and remote sensing. The developed CNN-LSTM approach improved its practicability and feasibility in order to forecast the Particulate Matter(PM2.5) concentration was also verified in the model. The DNN structure was developed that integrated LSTM and CNN based on the historical data such as cumulated wind speed, duration of rain, and concentration of PM 2.5The latest research in this area recommended that CNN could explore more spatial features and LSTM can reveal phonological features, which together play a significant role in crop yield prediction. However, the method employed histogram-based tensor alteration fused different remote sensing data which combined multisource data with a various resolution for feature extraction remained challenging,

Bondre and Mahagonkar [14] utilized ML techniques to predict the crop yield and manure recommendation. The yield prediction was a major issue in agriculture which was overcome by developing a machine learning algorithm. The performance of the developed model was evaluated for estimating crop production in agriculture. An advantage of the developed model was that earlier data was utilized for crop prediction and by applying ML algorithms like random forest and SVM the data also recommended a suitable fertilizer for every particular crop. However, the smart irrigation system for farms to get a higher yield method was not implemented.

Devika and Ananthi [15] utilized data mining techniques to predict the annual yield of major crops. Farmers were opposed to harvesting the yield because of insufficient availability of water sources and unpredictable weather variations but these issues were overcome by developing a data mining method. The developed model was gathering crop growing documents that used to be stored and analyzed for valuable crop yield prediction. In some of the data mining actions, the training data can be collected from the previous documents and the gathered documents were used in the phase of training which

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

has to exploit. An advantage of the developed model was that the highest level of crop yield prediction was obtained only in sugarcane, cotton.

ADVANTAGES

- This study includes the comparison of various previous methodologies proposed using different datasets and with different characteristics and accomplishments offerings and ingredients or not.

LITERATURE SURVEY:

Author:M Rashid methodology:

than 95.6\% of detection accuracy and 0.048 s of classification time cost. We transform the packed malware variants detection problem to a system calls classification problem. To reduce the obfuscation which is caused by packers, we first extract sensitive system calls and abandon obfuscated system calls. Then we organize these sensitive system calls as a vector which will be sent to our neural net- works later. As system call is a coarse-gained and sparse representation of executables, it causes bad training approximation and feature generalization. So we next propose our principal component initialized multi-layers neural networks to efficiently and effectively train and detect malicious instance with these sparse vectors. Our approach contains the following two phases, a training phase and a detection phase. The work shown, in training phase, we monitor the system interactions of executables

in Cuckoo sandbox to obtain the system calls. Each profile of executables we got from Cuckoo sandbox contains several fields: time-stamp, system call, base address, file name, executing times, etc. We only consider system calls since it can give us enough information to describe characteristics of behaviors of malware while reducing the noise and redundant.

Advantages:

Overcome the effect of unpacking behaviors of packers which add noisy information to the real behaviors of executables, which has a bad effect on accuracy.

Disadvantages:

It might be attacked by adversaries which causes security problem

Year: 2019

Author: D Elavarasan

Methodology:

clearly show that our proposed method is able to outperform all the other

which further boosts performance. Lastly, we implement the ensemble of ELMs in parallel

using GPUs to perform intrusion detection in real time.

Advantages:

and gives state-of-the-art performance for intrusion detection

Disadvantages:

The system cannot determine the new type of attack if it is not trained on it.

for a time interval, PV is the variance of payload packet length for a time

This feature is depending on the rambling packet payload length in the traffic flow. Our presented results and experiment produced by this research are significant and encourage other researchers and us to expand the work as future work.

Advantages:

components against intruders.

Disadvantages:

networks will harm the speed required.

Year: 2017

Author: Mozammel Chowdhury, Azizur Rahman

Methodology:

Exfiltration of sensitive data by malicious software or malware is a serious cyber threat around the world that has catastrophic effect on businesses, research organizations, national intelligence, as well as individuals. Thousands of cyber criminals attempt every day to attack computer systems by employing malicious software with an intention to breach crucial data, damage or manipulate data, or to make illegal financial transfers. Protection of this data is therefore, a critical concern in the research community. We propose a comprehensive framework to classify and detect malicious software to protect sensitive data against malicious threats using data mining and machine learning classification techniques. A hybrid framework is used for malware classification integrating a binary associative memory (BAM) with a multilayer perceptron (MLP) neural network by using both signature-based and behavior-based features analysis. In this work, we employ signature- based n-gram features and behavior-based API (Application Programming Interface) call sequences for malware analysis In this work, we employ a robust and efficient approach for malware classification and detection by analyzing both signature-based and anomaly-based features. Experimental results confirm the superiority of the proposed approach over other similar

methods. The proposed scheme for malware classification and detection is consisted of the following major components: (i) Pre-processing, (ii) Features extraction, (iii) Feature refinement/selection, (v) Classification, and (vi) Detection. Classification process is divided into two stages: training and testing. In the training phase, a training set of malicious and benign files is provided to the system. The learning algorithm trains a classifier. The classifier learns from the labeled data samples. In the testing phase, a set of new malicious and benign files are fed into the classifier and classified as malware or cleanware. In this work, we propose a hybrid framework for malware classification integrating a binary associative memory (BAM) with a multilayer perceptron (MLP) neural network. This is a robust and efficient approach for malware classification and detection using a hybrid framework with combination of a binary associative memory (BAM) and a multilayer perceptron (MLP) neural network. The BAM network can significantly reduce feature dimensions collected from a large malware dataset. We employ hybrid features for malware analysis by integrating both signature-based and behavior-based features that clearly increases classification and detection accuracy.

Advantages:

The BAM network can significantly reduce feature dimensions collected from a large malware dataset

Disadvantages:

False positive is high

Title: A flexible and extensible framework for agricultural Crop Yield Prediction.

Year: 2020

Author: Thai Le, Suhang Wang

Methodology:

Circulation of fake news, i.e., false or misleading pieces of information, on social media is not only detrimental to individuals' knowledge but is also creating an erosion of trust in society. Fake news has been promoted with deliberate intention to widen political divides, to undermine citizens' confidence in public figures, and even to create confusion and doubts among communities. Hence, any quantity of fake news is intolerable and should be carefully examined and combated. Due to the high-stakes of fake news detection in practice, therefore, tremendous efforts have been taken to develop fake news detection models that can auto-

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

detect fake news with high accuracies. In an attempt to solve these challenges, we propose MALCOM, a novel framework that can generate realistic and relevant comments in an end- to-end fashion to attack fake news detection models, that works for both black box and white box attacks. The main contributions are: This is the first work proposing an attack model against neural fake news detectors, in which adversaries can post malicious comments toward news articles to mislead cutting edge fake news detectors. Different from prior adversarial literature, our work generates adversarial texts (e.g., comments, replies) with high quality and relevancy at the sentence level in an end-to-end fashion (instead of the manipulation at the character or word level). Our model can fool five top-notch neural fake news detectors to always output real news and fake news 94% and 93.5% of the time on average. Moreover, our model can mislead black-box classifiers to always output real news 90% of the time on

average. We also compare our attack model with four baselines across two real-world datasets, not only on attack performance but also on generated quality, coherency, transferability, and robustness

Advantages:

Malcom is shown to be more robust even under the condition when a rigorous defense system works against malicious comments

Disadvantages:

Whether or not comments generated using one sub-domain can be transferable to another is also out of scope of this model

SYSTEM REQUIREMENTS

HARDWARE REQUIREMENTS

- Processor - AMD RYZEN–7
- RAM - 4GB (min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

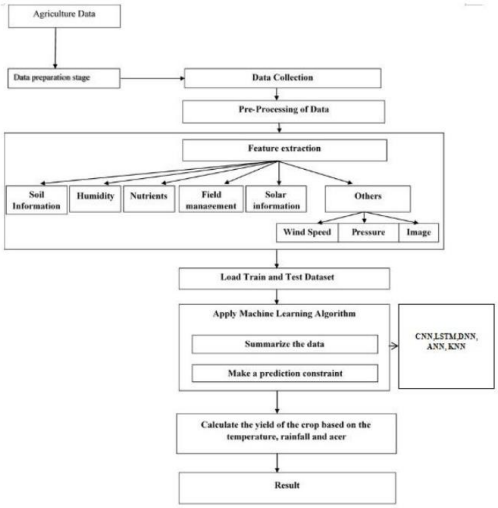
SOFTWARE REQUIREMENTS

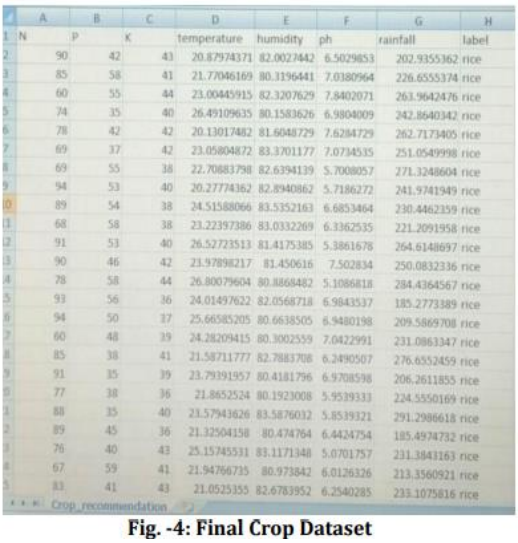
- Operating system - Windows 7 Ultimate.
- Coding Language - python
- Frontend - python
- Back-End - JSP
- Designing - Html, CSS, JS.

SYSTEM DESIGN:

SYSTEM ARCHITECTURE:

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362





	N	P	K	temperature	humidity	ph	rainfall	label
2	90	42	43	20.87974371	82.0027442	6.5029853	202.9355362	rice
3	85	58	41	21.77046169	80.3196441	7.0380964	226.6555374	rice
4	60	55	44	23.00445915	82.3207629	7.8402071	263.9642476	rice
5	74	35	40	26.49109635	80.1583626	6.9804009	242.8640342	rice
6	78	42	42	20.13017482	81.6048729	7.6284729	262.7173405	rice
7	69	37	42	23.05804872	83.3701177	7.0734535	251.0549998	rice
8	69	55	38	22.70883798	82.6394139	5.7008057	271.3248604	rice
9	94	53	40	20.27774362	82.8940862	5.7186272	241.9741949	rice
10	89	54	38	24.51588066	83.5352163	6.6853464	230.4462359	rice
11	68	58	38	23.22397386	83.0332269	6.3362535	221.2091958	rice
12	91	53	40	26.52723513	81.4175385	5.3861678	264.6148697	rice
13	90	46	42	23.97898217	81.450616	7.502834	250.0832336	rice
14	78	58	44	26.80079604	80.8868482	5.1086818	284.4364567	rice
15	93	56	36	24.01497622	82.0568718	6.9843537	185.2773389	rice
16	94	50	37	25.66585205	80.6638505	6.9480198	209.5869708	rice
17	60	48	39	24.28209415	80.3002559	7.0422991	231.0863347	rice
18	85	38	41	21.58711777	82.7883708	6.2490507	276.6552459	rice
19	91	35	39	23.79391957	80.4181796	6.9708598	206.2611855	rice
20	77	38	36	21.8652524	80.1923008	5.9539333	224.5550169	rice
21	88	35	40	23.57943626	83.5876032	5.8539321	291.2986618	rice
22	89	45	36	21.32504158	80.474764	6.4424754	185.4974732	rice
23	76	40	43	25.15745531	83.1171348	5.0701757	231.3843163	rice
24	67	59	41	21.94766735	80.973842	6.0126326	213.3560921	rice
25	83	41	43	21.0525355	82.6783952	6.2540285	233.1075816	rice

Crop_recommendation

Fig. -4: Final Crop Dataset

METHODOLOGY:

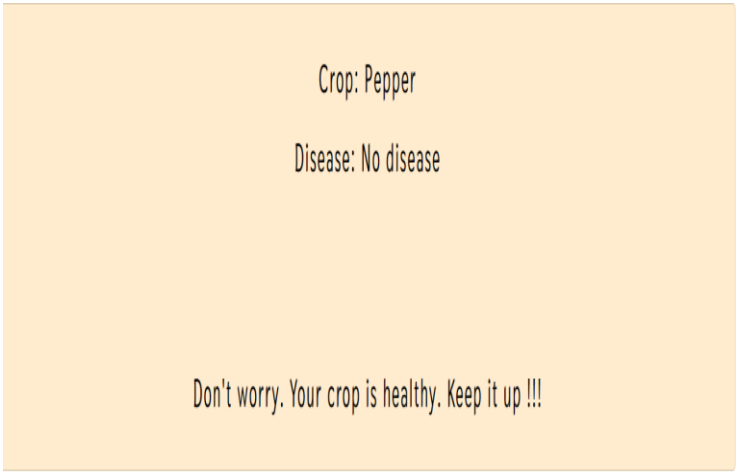
The methods and the processes to be carried out for implementation of our predictive system is stated below in a sequential manner.

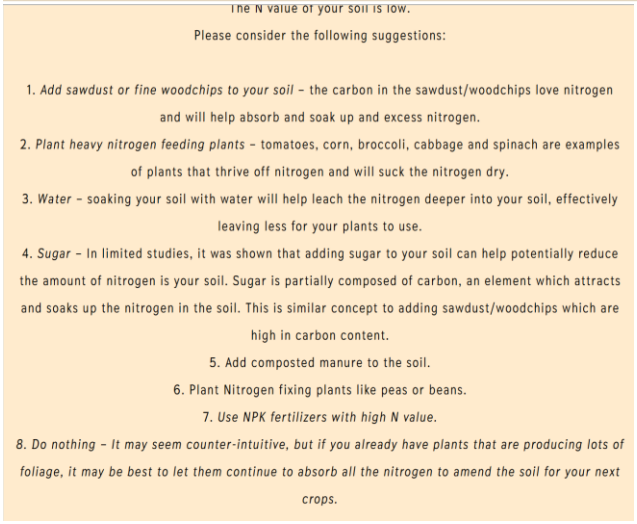
4.1 Dataset Collection

The first step that we perform during machine learning project development is the collection of dataset. The dataset we obtain from various platforms is the raw data having a tremendous amount of errors and ambiguities in it. In this project we have obtained our data from an open source platform known to as kaggle. Our dataset is basically a collection of two more intricate datasets. The two datasets are :

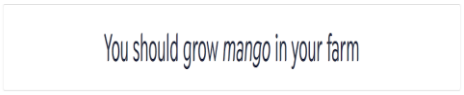
soil content dataset (consisting information about ratios of Nitrogen(N), Phosphorous(P), and Potassium(K) and ph of the soil) and climatic condition dataset (containing information about rainfall, humidity and temperature).The final dataset comprises about of about 2200 rows and about 8 columns. “Fig. 4” depicts our final dataset.

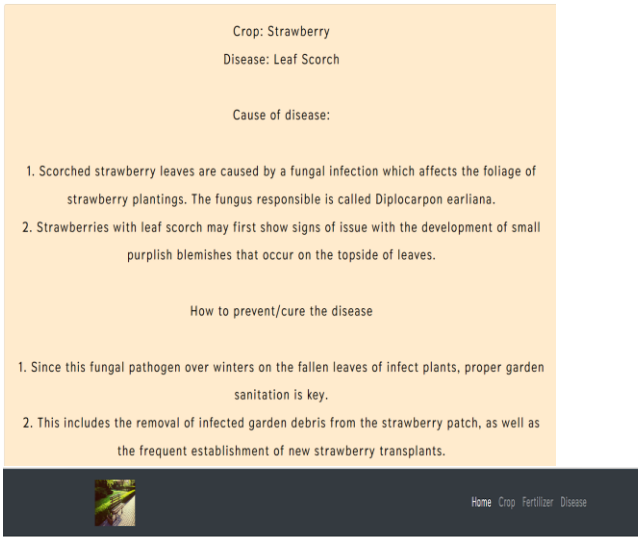
OUTPUT SCREENS:

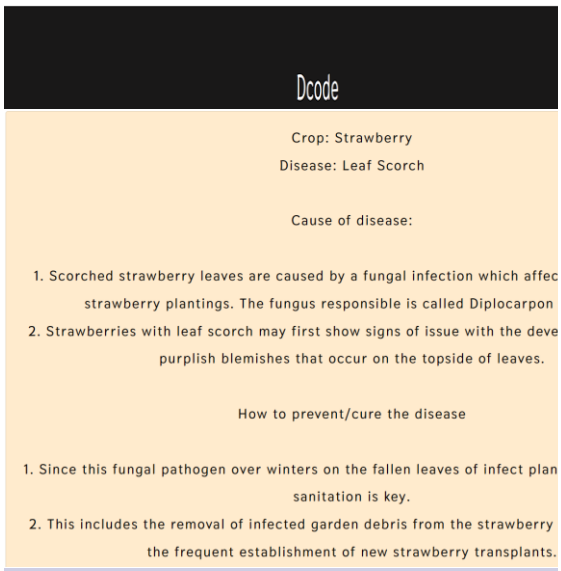


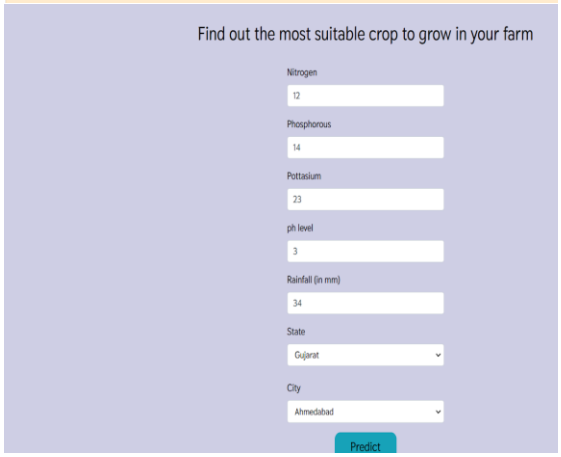


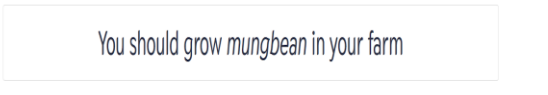
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

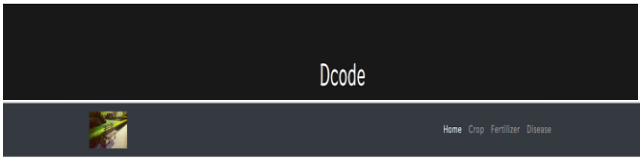


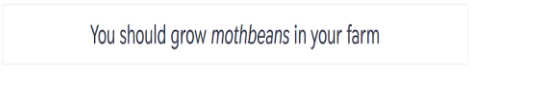












CONCLUSION:

The present research work discussed about the variety of features that are mainly dependent on the data availability and each of the research will investigated CYP using ML algorithms

that differed from the features. The features were chosen based upon the geological position, scale, and crop features and these choices were mainly dependent upon the data-set availability, but the more features usage was not always giving better results. Therefore, finding the fewer best performing features were tested that also have been utilized for the studies. Most of the exiting models utilized Neural networks, random forests, KNN regression techniques for CYP and a variety of ML techniques were also used for best prediction. From the studies most of the common algorithms used were CNN, LSTM, DNN algorithms but still improvement was still required further in CYP. The present research shows several existing models that consider elements such as temperature, weather condition, performing models for the effective crop yield prediction. Ultimately, the experimental study showed the combination of ML with the agricultural domain field for improving the advancement in crop prediction. However, still more improvement in feature selection was required in terms of temperature variation aspects effects on agriculture. In the further studies, the key possibility that should be concentrated such as firstly the delay to border topographical areas required additional-explicit treatment. Next, a non- parametric portion of the model using machine learning algorithm and thirdly, using features from deterministic crop models to get perfect statistical CO_2 fertilization.

FUTURE SCOPE:

In future research, additional behaviors , a non- parametric portion of the model using machine learning algorithm and thirdly, using features from deterministic crop models to get perfect statistical CO_2 fertilization. By following above-mentioned objectives, the crop yield estimation would be improved by further researchers. Additionally, in the crop yield estimation, fertilizer should also be considered for executing soil forecasts that agriculturalist to make a better judgment based on the situation of low crop yield estimation. Based on the outcomes obtained for the study further we need to build and develop a model based on DL for CYP.

REFERENCES:

[1] https://en.wikipedia.org/wiki/Agriculture_in_India. [2] https://www.kaggle.com/datasets/atharvaingle/croprecommendation-dataset [3] Crop Recommendation System for Precision Agriculture S.Pudumalar*, E.Ramanujam*, R.Harine Rajashree, C.Kavya, T.Kiruthika, J.Nisha. In 8th International Conference on Advanced Computing (ICoAC), IEEE, 2017. [4] Prediction of Crop Cultivation – Rikhsit K. Solanki, D Bein, J. A. Vasko & N. Rale. In 9th Annual Computing and communication Workshop and Conference (CCWC), IEEE, 2019. [5] In the 5th International Conference on Communication and Electronics Systems (ICCES), June 2020. Dr. J. N. Kumar, V. Spandana, V.S. Vaishnavi have presented their research on Supervised Machine learning Approach for Crop Yield Prediction in Agricultural Sector. [6] Improving Crop Productivity Through A Crop Recommendation System Using Ensembling Technique – N. K. Cauvey, Nidhi H. Kulkarni, Prof. B. Sagar, N. In 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), pp. 114-119. IEEE, 2018. [7] Nilesh Dumber, Omkar Chikane, Gitesh Moore, "System for Agriculture Recommendation using Data Mining" , EISSN : 2454- 9916 | Volume : 1 | Issue : 4 | NOV 2014. [8] Tom M. Mitchell, Machine Learning, India Edition 2013, McGraw Hill Education. [9] David C.

Rose, William J. Sutherland, Caroline Parker, "Decision support tools for agriculture: Towards effective design and delivery", Agricultural Systems 149 (2016) 165–174, journal homepage: www.elsevier.com/ locate/agsy.