NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY
(AN AUTONOMOUS INSTITUTION, AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY,
BELGAUM, APPROVED BY AICTE & GOVT.OF KARNATAKA

PROJECT REPORT – PHASE 0

on

CROP RECOMMENDER SYSTEM USING MACHINE LEARNING

Submitted in partial fulfilment of the requirement for the award of Degree of

Bachelor of Engineering

in

Computer Science and Engineering

Submitted by:

Amreshwar Singh	1NT19CS026
Aman Kumar	1NT19CS024
Aryan Burman	1NT19CS039
Mohammed Abrar	1NT19CS220

Under the Guidance of

Dr. Saroja Devi H
HOD,Professor,Dept. of CS&E, NMIT

Department of Computer Science and Engineering
(Accredited by NBA Tier-1)

2021-2022

25/01/2022

NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY

(AN AUTONOMOUS INSTITUTION, AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM
, APPROVED BY AICTE & GOVT.OF KARNATAKA)

Department of Computer Science and Engineering

(Accredited by NBA Tier-1)

CERTIFICATE

This is to certify that the **Crop Recommender System Using Machine Learning** is an authentic work carried out by **Amreshwar Singh (1NT19CS026) ,Aman Kumar (1NT19CS024) , Aryan Burman (1NT19CS039) and Mohammed Abrar (1NT19CS220)** bonafide students of **Nitte Meenakshi Institute of Technology**, Bangalore in partial fulfilment for the award of the degree of ***Bachelor of Engineering*** in COMPUTER SCIENCE AND ENGINEERING of Visvesvaraya Technological University, Belgavi during the academic year ***2021-2022.*** It is certified that all corrections and suggestions indicated during the internal assessment has been incorporated in the report. This project has been approved as it satisfies the academic requirement in respect of project work presented for the said degree.

Internal Guide	**Signature of the HOD**	**Signature of Principal**
________________	________________	________________
Dr. Sarojadevi H.	Dr. Sarojadevi H.	Dr. H. C. Nagaraj
HOD ,Professor , Dept. CSE,	Professor, Head, Dept. CSE,	Principal,
NMIT Bangalore	NMIT Bangalore	NMIT,
		Bangalore

Signature of Examiner (Internals**)**	**Signature of Examiner (**SEE**)**
1. ____________________	1. ____________________
2. ____________________	2. ____________________

25/01/2022

DECLARATION

We hereby declare that

(i) The project work is our original work

(ii) This Project work has not been submitted for the award of any degree or examination at any other university/College/Institute.

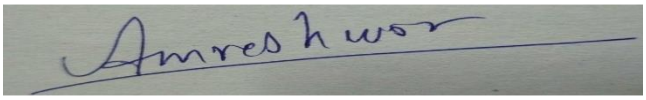
(iii) This Project Work does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

(iv) This Project Work does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been re-written but the general information attributed to them has been referenced;

b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced.

(v) This Project Work does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

NAME	USN	Signature
Amreshwar Singh	1NT19CS026	
Aman Kumar	1NT19CS024	
Aryan Burman	1NT19CS039	
Mohammed Abrar	1NT19CS220	

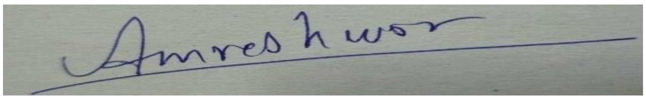
ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible, whose constant guidance and encouragement crowned our effort with success. I express my sincere gratitude to our Principal **Dr. H. C. Nagaraj**, Nitte Meenakshi Institute of Technology for providing facilities.

We wish to thank our HoD**, Dr. Sarojadevi H.** for the excellent environment created to further educational growth in our college. We also thank him for the invaluable guidance provided which has helped in the creation of a better project.

I hereby like to thank our **Dr. Sarojadevi H ,HOD, Professor** , Department of Computer Science & Engineering on her periodic inspection, time to time evaluation of the project and help to bring the project to the present form.

Thanks to our Departmental Project coordinators. We also thank all our friends, teaching and non-teaching staff at NMIT, Bangalore, for all the direct and indirect help provided in the completion of the project.

NAME	**USN**	**Signature**
Amreshwar Singh	1NT19CS026	
Aman Kumar	1NT19CS024	
Aryan Burman	1NT19CS039	
Mohammed Abrar	1NT19CS220	

ABSTRACT

Agriculture and its united areas are without a doubt the biggest suppliers of vocations in rustic India. The horticulture area is likewise a critical giver element to the nation's Gross Domestic Product (GDP). Gift to the nation is the staggering size of the farming area. Be that as it may, deplorable is the yield per hectare of harvests in contrast with worldwide principles. This is one of the potential reasons for a higher self destruction rate among minimal ranchers in India. This paper proposes a suitable and easy to understand yield forecast framework for the ranchers. The proposed framework gives availability to ranchers by means of a versatile application. GPS assists with distinguishing the client area. The client gives the region and soil type as info. AI calculations permit picking the most productive harvest list or foreseeing the yield for a client chose crop. To foresee the harvest yield, chose Machine Learning calculations like Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbor (KNN) are utilized. Among them, the Random Forest showed the best outcomes with 95% exactness. Furthermore, the framework likewise recommends the best an ideal opportunity to utilize the manures to support up the yield.

TABLE OF CONTENTS

List of figures:-

INTRODUCTION

The main occupation rehearsed in our nation is farming. The economy of India would have fallen assuming there was no agribusiness. In the public GDP of India, around 15-20% is contributed by the farming area [1]. Farming assumes a vital part in generally speaking improvement just as in the monetary area of the country. Practically 60% of the land in our nation is used for horticulture to fulfill the requirements of individuals in our country. In the worldwide economy, horticulture assumes a significant part. With the proceeded with development of the human populace, strain will increment on the agribusiness framework. Agribusiness, presently likewise called advanced horticulture, has risen up out of the new logical fields which utilize information escalated ways to deal with advance ranch usefulness while limiting ecological impacts [2]. Data accumulated in cutting edge crop development depend on various kinds of sensors which permit a more exact and speedy choices to be taken, upgrading working climate understanding (cooperation between adaptable harvest, soil, and climate).

AI (ML) has arisen, in blend with huge information innovation and elite execution PC innovation, to set out new open doors for finding, evaluating and involving information concentrated farming cycles. Our ranchers genuinely should create gains involving logical methodologies and advancements in farming. In India, wheat and paddy are the main food crops. Consistently ranchers in India take tough choices that influence their salaries. Utilizing new procedures and advanced utilizations of AI ranchers can anticipate the harvest's cost, can know the compost for the dirt, and plan the best yield for that season. Composts improves the dirt wellbeing as well as it builds holding limit of water. In this way, knowing appropriate utilization of manure helps in getting more yield. Reaping crops where climatic conditions are not appropriate for the development of harvest prompts a decrease in yield. Henceforth appropriate yield ought to be chosen to develop as indicated by the conditions. And furthermore having a thought of patterns and cost of the harvests in forthcoming years will bring about getting more benefits for the ranchers.

In the current circumstance, numerous ranchers need more information regarding how to utilize innovation to develop crops and furthermore, they don't know about the advantages they get from cultivating. Likewise, cultivating can be expanded by understanding the yield's presentation in the specific land and different ecological conditions. These are a portion of the difficulties looked by the ranchers consistently. In creating Machine Learning Models, coming up next are the principle challenges. Preparing information overfitting. Intricacy of the information pipeline. The accuracy and interpretability of the model equilibrium. Assortment and capacity of huge amounts of homestead datasets. This paper proposes three unique models which are productive for crop the board. The primary model predicts the compost to be utilized in view of soil supplements and natural conditions. The subsequent models foresee the yields which generally plausible to fill in the given climatic conditions. The third model predicts the future cost of the yield from the past costs and patterns. Additionally, legitimate Data Visualization strategies are utilized all the above proposed models.

1.1 Background

In Indian economy and work farming assumes significant part. The most widely recognized issue looked by the Indian ranchers is they don't select yield in light of the need of soil, thus they face genuine misfortune in usefulness. This issue can be tended to through accuracy farming. This technique thinks about three boundaries, viz: soil qualities, soil types and harvest yield information assortment in light of these boundaries recommending the rancher appropriate harvest to be developed. Accuracy agribusiness helps in decrease of non appropriate harvest which for sure expands efficiency, aside from the accompanying benefits like viability in input just as result and better decision making for cultivating.

In India, Agriculture is vital and rehearsed extensively the nation over. This industry assumes a significant monetary part in the nation's turn of events. Thus, three principle models are should have been considered while developing harvests. Picking manures, choosing appropriate harvests as indicated by the district's climatic condition and Knowledge of yield cost are the standards that benefit the ranchers and furthermore helps the monetary development of the country. This paper proposes three distinct models for every one of the previously mentioned rules. The principal model is the Fertilizer indicator, which predicts the reasonable compost that can be utilized for the given soil constituents. The subsequent model is the harvest indicator, which predicts the three no doubt growable yields in light of the given climatic conditions. The third model is the future harvest value indicator, which predicts the yield cost in future in light of past patterns and costs.

1.2 Brief history of Technology/concept

The historical backdrop of farming records the taming of plants and animals and the turn of events and spread of strategies for raising them beneficially. Horticulture started autonomously in various pieces of the globe, and incorporated a different scope of taxa. Something like eleven separate districts of the Old and New World were involved as free focuses of beginning.

Wild grains were gathered and eaten from somewhere around 105,000 years prior. Notwithstanding, taming didn't happen until some other time. Beginning from around 9500 BC, the eight Neolithic author crops - emmer wheat, einkorn wheat, hulled grain, peas, lentils, harsh vetch, chickpeas, and flax - were developed in the Levant. Rye might have been developed before, however this guarantee stays disputable. Rice was tamed in China by 6200 BC with earliest known development from 5700 BC, trailed by mung, soy and azuki beans. Pigs were trained in Mesopotamia around 11,000 years prior, trailed by sheep between 11,000 years prior and 9000 BC. Dairy cattle were tamed from the wild aurochs in the space of present day Turkey and India around 8500 BC. Sorghum was trained in the Sahel district of Africa by 3000 BC. Camels were trained late, maybe around 3000 BC.

1.3 Applications

Plants and harvests infection recognition: By 2050, human horticultural harvest yield should increment by an expected 70 percent to support the normal populace size. By a specific measurements, crop sicknesses right now lessen the yield of the six most significant food crops by 42%, and a few ranches are cleared out completely on a yearly premise. In this manner, it happens to most extreme significance to track down techniques by utilizing innovation for precise yield sickness location. This is the place where AI strategies can help. Profound learning calculations can be prepared on yields and plants pictures with great precision for crop sickness identification. The test is to get the plant/crop pictures and afterward named pictures. One of the common methods utilized presently is Unmanned Aerial Vehicles (UAVs) combined with huge scope backend frameworks including AI models for distinguishing crop sicknesses. To address the test related with information assortment, displaying procedure, for example, generative antagonistic organization (GAN) can be utilized to produce engineered information utilizing the harvest infection pictures. One of the other test for preparing models having high exactness is class awkwardness in the gathered information. This is the place where DC-GAN (Deep convolutional GAN) assumes an incredible part to mitigate the class irregularity issue by creating engineered pictures. A profound convolutional neural organization (CNN) model (like EfficientNet) can then be prepared to group and distinguish crops/plants sicknesses. The CNN model can be prepared to recognize afflictions that made their actual presence on the leaf as well as stem of the yield and distinguish sicknesses. Review that generative ill-disposed organizations are sets of neural organizations that are isolated into two jobs: generator and discriminator. The generator figures out how to foster manufactured pictures of some class, while the discriminator figures out how to perceive among genuine and engineered pictures. The models train off of one another to further develop results.

Crop yield expectation: Predicting the yield precisely will assist ranchers with knowing when they should begin reaping so they can boost their benefits by selling it at a fitting cost. Crop yield expectation is tied in with estimating the normal yield of farming harvests in a given period. Crop yield forecast is amazingly difficult because of its reliance on various factors, for example, crop genotype, natural elements, the board rehearses, and their associations. The AI models are worked to foresee crop yields by thinking about various variables that influence it like climate information (temperature, precipitation), soil dampness sensors, cosmology pictures and so forth, subsequently anticipating precise yield esteems for an agribusiness field before collect time. These procedures can be utilized by ranchers on regular schedule with high precision which empower them to settle on choices on when to reap crops, how much pesticide should be applied and what composts should be utilized. Profound learning models can be utilized to anticipate farming creation in enormous scope with a precise assessment of the yield. This will assist ranchers with settling on significant choices connected with editing examples and yield the board prompting better yields during harvest season. Calculations, for example, multi-straight relapse, Lasso relapse, LightGBM, irregular backwoods, XGBoost and profound neural organizations (CNN, LSTM and so on) have been utilized for crop yield expectations.

Crop column identification: Crop line discovery is a critical component in creating vision based route robots in rural advanced mechanics. Ongoing work on crop line discovery has utilized profound learning based techniques along these lines conquering the significant difficulties in carrying out a certifiable vision based route framework. A portion of the vital viewpoints in crop column recognition are weed thickness, development stages, shadows and discontinuities and so on CNN models can be utilized for crop column location in light of weed thickness, development stages, shadows and so on

Recognize water pressure in plant: Plant water pressure might happen because of the restricted accessibility of water to the roots/soil or because of expanded happening. These elements antagonistically influence plant physiology and photosynthetic capacity to the degree that it has been displayed to have inhibitory impacts in both development and yield. Early distinguishing proof of plant water pressure status empowers reasonable remedial measures to be applied to acquire the normal harvest yield. It is important to distinguish potential plant water pressure during the beginning phases of development to present remedial water system and mitigate pressure. This is the place where AI strategies come into picture. AI methods can be utilized to appraise leaf water content (LWC) which is then additionally used to gauge water pressure in the plants. Leaf water content (LWC) is an action that can be utilized to assess water content and distinguish focused on plants. LWC during the early harvest development stages is a significant mark of plant efficiency and yield. Various strategies can be utilized for information assortment. They incorporate use of sensors or UAVs. Use of sensors can, nonetheless, end up being pricey. Gathering and regressor techniques can be utilized to anticipate the LWC esteem. Furthermore, arrangement models can be utilized to order the water pressure in view of LWC and different boundaries.

Crop planning: Crop type planning at the field level is basic for an assortment of uses in rural observing. Crop type planning at the field goal is an essential to planning ranch the executives and yield results at large spatial scale. This undertaking is even more critical in when populaces in food-uncertain districts proceed to increment and environmental change is anticipated to unfavorably influence worldwide horticulture. Customarily, crop type data has been acquired from field reviews and censuses, however such overviews are costly and tedious to lead. This is the place where AI methods are applied on the satellite information for crop type maps. Arrangement calculations, for example, LDA, irregular backwoods can be utilized for crop grouping and planning.

Crop choice expectation: Machine learning strategies can be utilized to assist the ranchers with choosing the harvest proficiently and expand crop yield with negligible expense. The AI models can be prepared to foresee most fitting harvest determination and yield for various areas. One will be needed to choose various sorts of yields, recognize elements and afterward train model to order the harvest choice for various districts. Calculations, for example, SVM, arbitrary woodland, strategic relapse, profound neural organizations and so forth, can be utilized to prepare such models. Highlights utilized in such models can be connected with climate boundaries (precipitation, temperature and so forth), composts utilized, land type, soil-related data and so on

Water system location: Detecting water system is basic to comprehend water use and advance better water the board. Such information will possibly empower the investigation of environmental change sway on horticultural water sources, screen water utilization, assist with recognizing water burglary and unlawful farming and illuminate strategy choices and guidelines connected with water consistence and the board. AI methods can be utilized for water system discovery. In any case, this is an intricate issue to settle with the assistance of ML methods because of the absence of organized and named information accessible that are revolved around water system frameworks. This is the place where pre-prepared models can help. This will be order models and the objective mark is a paired variable showing whether or not the land in the picture is forever flooded. CNN models can be prepared to group the land as watered or in any case.
Ground water level expectation: Groundwater is the biggest stockpiling of freshwater assets, which fills in as the significant stock for the greater part of the human utilization through farming, modern, and homegrown water supply. Profound neural organizations can be prepared to estimate ground-water levels. Profound learning techniques are known to deliver precise outcomes even with the restricted data accessible for this situation, which is for the most part satellite information and hydro-meteorological boundaries.

1.4 Research motivation and Problem statement

1.4.1 Research motivation

India is probably the greatest maker of horticultural items and still has exceptionally less homestead usefulness. Efficiency should be expanded so ranchers can get additional compensation from a similar land parcel with less work. Accuracy farming gives a method for getting it done.

Besides, the experience of executing appraisal practice propels in an inexorably asset obliged climate, in certain examples moving toward the asset limitations of altogether more unfortunate states, delivers the examination discoveries possibly valuable to other Indian Technological confronting comparative difficulties.

1.4.2 Problem Statement

- The goal of the paper is to assemble a recommender framework to expand the harvest yield
- The framework helps the ranchers in choosing an appropriate yield for their farming area in light of the necessary boundaries.
- TheThe proposed recommender model is planned, prepared and tried to prescribe an appropriate yield to the rancher.

1.5 Research objectives and contributions

1.5.1 Primary objectives

This project aims to help the farmers to cultivate the best crop, predict the price of wheat, rice, Maize, etc, and also it will help to do the fertilizer analysis.

1.5.2 Main contributions

Through this project, we aim to optimize the current farming system by using Machine learning and A.I technology which will be implemented using python; a secure, performance and scalable language. This project will contribute in making an effective system of analysis in farming. The project will be very reliable for physicians, doctors, patients, etc. in the medical field. Patient records will be thorough and secure. The details of patient visits from clinics to hospitals will be updated and available on the blocks. Hence, the quality of diagnosis and consulting will improve abundantly.

Through this project, we aim to increase the production of crop. This study motivated to address some of the key challenges of farm-level crop yield estimation using optical sensors.

1.6 Organization of the report

We first explain briefly the Machine Learning data structure. Then we present issues in the present farming system and followed by the solutions. Continuing, we discuss the various advantages of modern technology. The last part of this report describes our model of a machine model based solution for farming and trust management which provide a solution to overcome the traditional management.

1.7 Summary

The number of articles included in this review was 40 in total. Twenty-five (25) of the presented articles were published in the journal «Computer and Electronics in Agriculture», six were published in the journal of «Biosystems Engineering», and the rest of the articles were published to the following journals: «Sensors», «Sustainability», «Real-Time Imagining», «Precision Agriculture», «Earth Observations and Remote Sensing», «Saudi Journal of Biological Sciences», «Scientific Reports», and «Computers in Industry». Among the articles, eight of them are related to applications of ML in livestock management, four articles are related to applications of ML in water management, four are related to soil management, while the largest number of them (i.e., 24 articles) are related to applications of ML in crop management.

From the investigation of these articles, it was observed that eight ML models have been carried out altogether. All the more explicitly, five ML models were executed in the methodologies on crop the executives, where the most well known models were ANNs (with most continuous harvest nearby wheat). In animals the board class, four ML models were executed, with most famous models being SVMs (most incessant animals type close by dairy cattle). For water the executives specifically evapotranspiration assessment, two ML models were carried out and the most often carried out were ANNs. At last, in the dirt administration classification, four ML models were executed, with the most well known one again being the ANN model.

LITERATURE SURVEY

i)Crop Management System Using Machine Learning

Introduction

Accuracy cultivating, as the name suggests, alludes to the applying of exact and appropriate all out of remark like manures, soil and so forth however as of late the pattern in farming has definitely advanced because of globalization .different variables have impacted the soundness of horticulture in India. numerous new innovations have been developed to recapture the wellbeing . One such strategy is accuracy agribusiness , at the legitimate opportunity to the gizzard for expanding its usefulness and expanding its yields But in farming it is critical that the suggestions made are exact and exact in light of the fact that in the event of blunders it might prompt weighty material and capital misfortune. suggestion of yields is one significant area in accuracy horticulture.

Related work

Dataset Collection: The dataset involving the dirt explicit properties which are gathered for Ramtek town tried at soil testing lab Nagpur, Maharashtra, India.
Crop Prediction utilizing outfit procedure: Ensemble is an information mining model otherwise called the Committee Methods or Model Combiners, That combiners, that consolidate the force of numerous models to gain more prominent forecast, effectiveness that any of its models could accomplish alone
It comprise of four machine students to be specific Naïve bayes,K-Nearest neighbor and CHAID and Random tree. The administrator relating to every student is situated.
On the off chance that ph is mid antacid AND profundity is over 100 AND water holding limit is LOW AND seepage is respectably well AND disintegration is moderate THEN PADDY

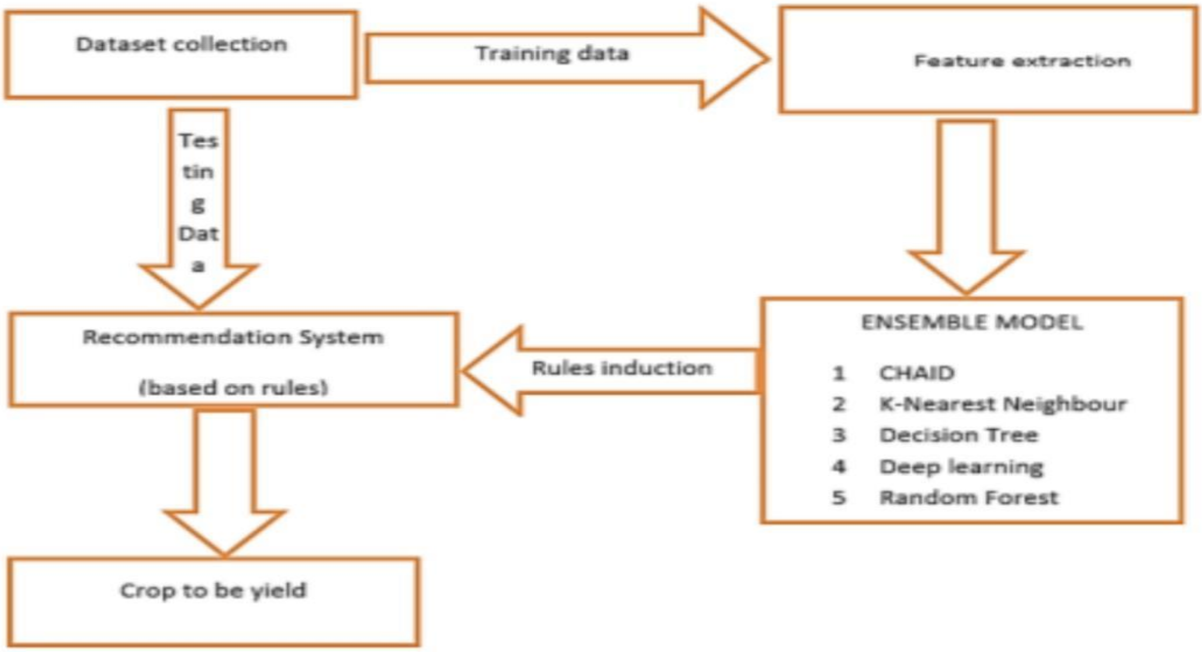


Fig 1. Proposed System Architecture.

fig – 1.1

Study of Tools and Technology

1. Unpredictable TREE: Random tree like that of a decision tree.But it fluctuates from subjective tree in a way that for each split simply a discretionary subset of attribute are open.

2. K-NEAREST NEIGHBOR: K-Nearest Neighbor can be used for both request and backslide. K-Nearest Neighbors is a non-complex computation which stores all of the open cases and classifiers new cased considering some closeness measure.

3. Subjective FOREST: Random boondocks square measure accomplice outfit learning approach for request, backslide and different endeavors

4. Decision TREE: Classifies data using the qualities. Tree involves decision centers and decision leafs. Centers can have somewhere around two branches which tends to the motivation for the characteristics attempted. Leafs centers makes a homogeneous result.

Summary:

Accuracy horticulture helps in decrease of non reasonable harvest which without a doubt builds usefulness, aside from the accompanying benefits like adequacy in input just as result and better decision making for cultivating. This strategy gives arrangements like proposing a suggestion framework through an outfit model with greater part casting a ballot methods utilizing irregular tree, CHAID, K _ Nearest Neighbor and Naive Bayes as student to suggest reasonable harvest in view of soil boundaries with high explicit exactness and proficiency. The grouped picture created by these methods comprises of ground truth measurable information and boundaries of it are climate, crop yield, state and region astute harvests to foresee the yield of a specific yield under specific climate condition.

ii)Advancing agricultural research using machine learning algorithms

Introduction

Increasing food demand will challenge the agricultural sector globally over the next decades. A sustainable solution to this challenge is to increase crop yield without massive cropland area expansion. Tis can be achieved by identifying and adopting best management practices. To do so requires a more detailed understanding of how crop yield is impacted by climate change2,3 and growing-season weather variability . Even with that knowledge, prediction is challenging because various factors interact with each other. For example, variability in soil type can interact with weather conditions and mitigate or aggravate climate-related impacts on crop yield . Additionally, seed genetics (G) and crop management decisions (M), interact with the effect of environment (E: soil and in-season weather conditions), thereby resulting in a near infinite number of combinations of G ×E ×M that can impact crop yield

Related work

1. Reducing the frequency of lowest vs. highest yields has been proposed as an effective means to increase food production in existing crop land

2. Weather data for maize (n=17,013) and soybean (n=24,848) involving US crop performance trials conducted in 28 states between 2016 to 2018 for maize and between 2014 to 2018 for soybean.(shown in the map)
3. For each crop, an algorithm was trained and used to estimate the yield.
4. The developed algorithms exhibited a high degree of accuracy when estimating yield in independent datasets .
5. Process-based models have been extensively used to evaluate the effect of weather and management on crop yield. However, to obtain accurate estimates, the models require extensive calibration, which is not a trivial task due to the large number of parameters.

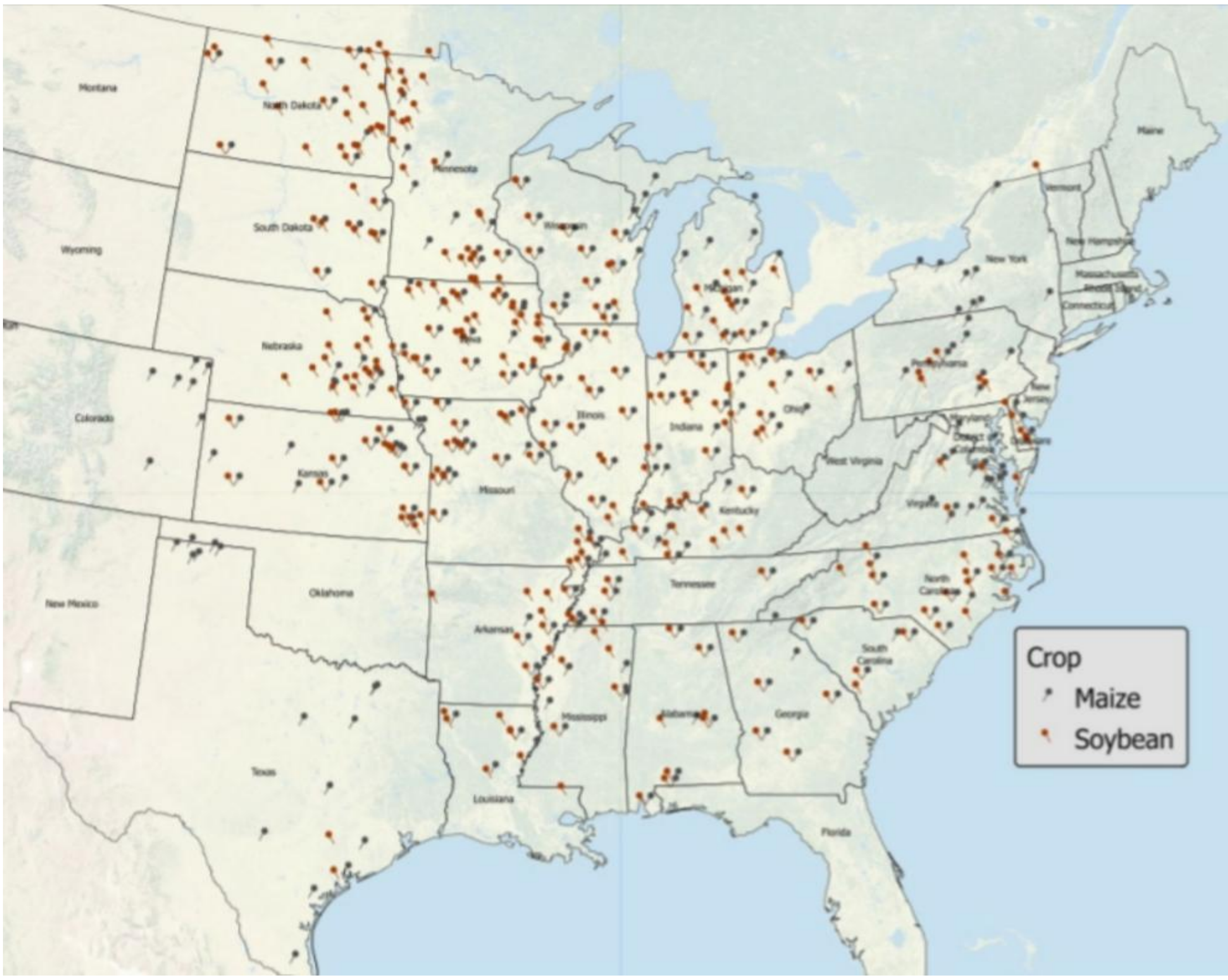


fig – 2.1

Study of Tools and Technology

1. Three management factors(G-seed genetic,E-soil and weather,M-management method) and their interactions are evaluated in a single location due to practical constraints (e.g., cost, logistics). By holding the background management constant, causal relationships are identified, and the effectiveness of the examined management practices is assessed.
2. **Extreme gradient boosting algorithm (XGBoost)** - is used estimate yield based on soil type and weather conditions (E), seed traits (G) and management practices (M) was developed.
3. Several Data Mining techniques were used in the algorithm like , the root mean square error (RMSE), mean absolute error (MAE) and mean centered yield.
4. Accumulated local effect plots for maize and soybean were recorded these included a)sowing date b)seeding rate c)Nitrogen fertilizer rate d)cumulative precipitation (mm) e)row spacing
5. **Producer self-reported data technique**, which can capture yield trends attributable to producer management choice across large regions, but such studies lack sufficient power relative to establishing causality and evaluating complex highorder G×E×M interactions.
6. Bayesian optimization was performed using “hyperopt” to efficiently tune the hypermeters.

Summary

Agrarian trials rehashed each year in many areas across the US create a tremendous measure of harvest yield and the executives datasets which are valuable for expansive inductions Such datasets have, until now, stayed disengaged from one another, and are hard to join, normalize, and appropriately break down. In the introduced work, we conquered these issues by growing enormous information bases and by utilizing the force of ML calculations. These calculations can progress farming exploration and help in uncovering a right now stowed away yield potential.

iii)SMART FARMING USING MACHINE LEARNING

Introduction

Agriculture is the backbone of the Indian economy. But agriculture in India is undergoing a structural change leading to a crisis situation. The relative contribution of agriculture to the GDP has been declining over time steadily. It is alarming that India is moving from being a self-reliant nation of food to a net importer of food. All these trends indicate that the agricultural sector in India is facing a crisis today
In the past farmers used to predict their yield from previous year yield experiences. Thus, for this kind of data analytics in crop prediction, there are different techniques or algorithms, and with the help of those algorithms, we can predict crop yield.

Therefore, this paper proposes an idea to predict the crop and yield of the crop based on the climatic conditions and historic data related to the crop .

Related work

1. To implement the system, we decided to focus on Maharashtra State only in India. Historical data about the crop and the climate at the district level was needed to implement the system.
2. This data has been gathered from the government website www.data.gov.in which includes State, District, Season, Crop, Area and Production.
3. The data about the climate conditions suitable for the particular crops has been collected from the Kaggle which includes Temperature, Humidity, Soil pH, Rainfall and class label is the Crop.
4. Exploratory Data Analysis: It refers to the critical process of performing initial investigations on data so as to discover patterns,to spot anomalies,to test hypotheses and to check assumptions with the help of summary statistics and graphical representations.

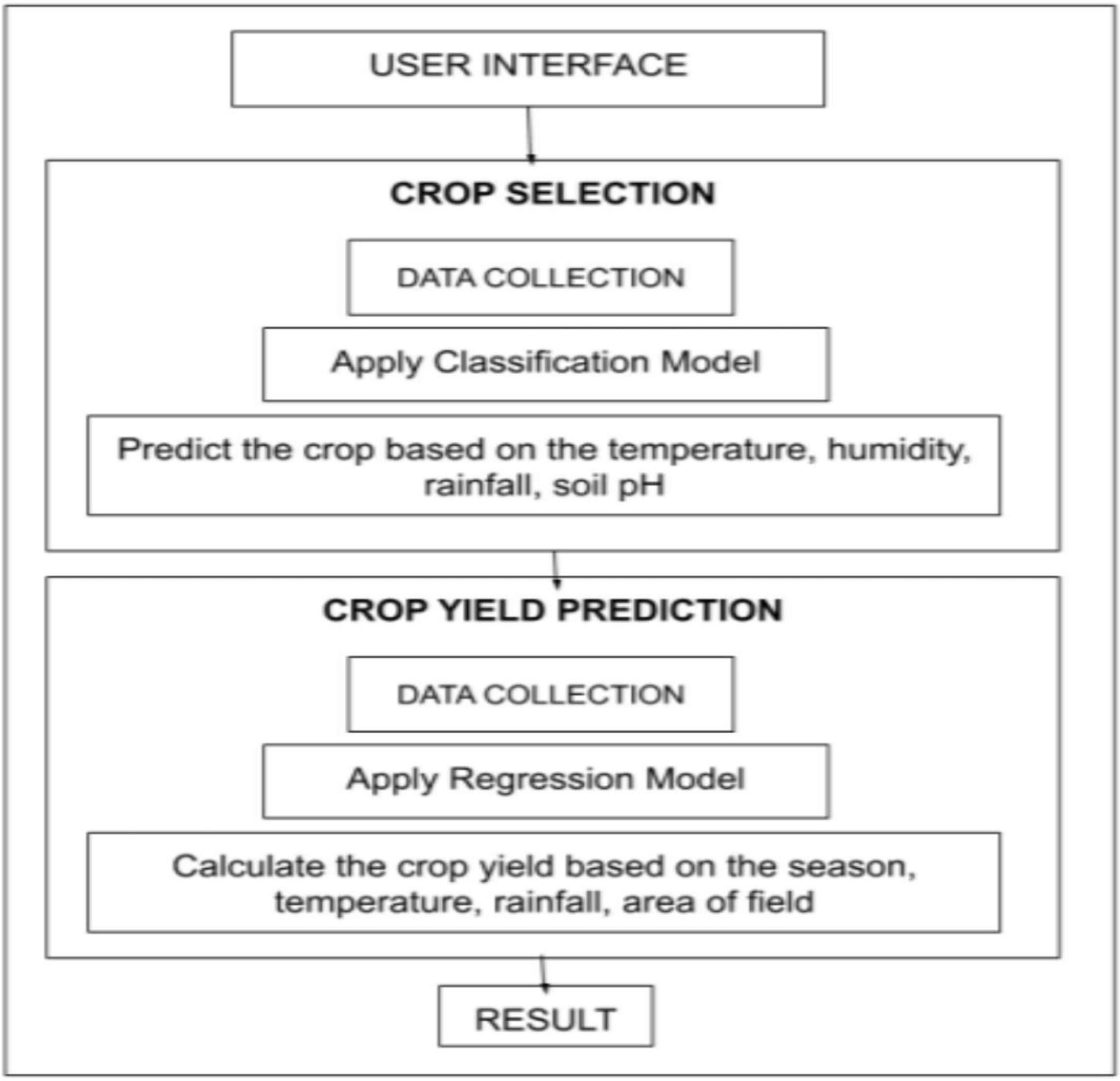


Fig. 3. Proposed Approach

Fig-3.1

Study of Tools and Technology

1. Data Cleaning: It is the process of preparing data for analysis by removing or modifying data that is incorrect, incomplete, irrelevant, duplicated, or improperly formatted.
2. Encoding: It is a required pre-processing step when working with categorical data for machine learning algorithms.
3. Feature Scaling: It is a technique to standardize the independent features present in the data in a fixed range. It is performed during the data pre-processing to handle highly varying magnitudes or values or units.
4. Data Partitioning: The Entire dataset is partitioned into 2 parts: for example, say, 75% of the dataset is used for training the model and 25% of the data is set aside to test the model

Summary

Crop and yield of the crop prediction using intelligent machine learning techniques may improve the crop planning decisions. For the Crop Prediction Module, the Cohen's Kappa score we got for the Naive Bayes Classification Model is about 95%. For the Crop Yield Prediction Module, the R-Squared value we got for the Random Forest Regression Model is more than 81%. Accurate forecasts of the climate parameters and better historic data of the crop would result in accurate crop and its yield forecast in the future. Also, the developed webpage is user friendly and can be made more informative by providing additional useful information like intercropping, fertilizers etc. to the user. We can create more interactive User Interface by adding chatbots and speech recognition systems

iv)Machine Learning in Agriculture

Introduction

In this paper, The works analyzed were categorized in (a) crop management, including applications on yield prediction, disease detection, weed detection crop quality, and species recognition; (b) livestock management, including applications on animal welfare and livestock production; (c) water management; and (d) soil management. The filtering and classification of the presented articles demonstrate how agriculture will benefit from machine learning technologies. By applying machine learning to sensor data, farm management systems are evolving into real time artificial intelligence enabled programs that provide rich recommendations and insights for farmer decision support and action but here we have only **taken crop,soil and water management**.

Related work

1. **Crop management/yield production** - A group of authors, developed a early yield mapping system for the identification of immature green citrus in a citrus grove under outdoor conditions. As all other relative studies, the aim of the study was to provide growers with yield-specific information to assist them to optimise their grove.
2. In another study, the authors developed a model for the estimation of grassland biomass (kg dry matter/ha/day) based on ANNs and multitemporal remote sensing data.
3. **Water Management-**In another study, the authors developed a computational method for the estimation of monthly mean evapotranspiration for arid and semi-arid regions.
4. In another study, a model was presented for the prediction of daily dew point temperature, based on ML.
5. **Soil Management-** In a study, it presented a method for the evaluation of soil drying for agricultural planning. The method accurately evaluates the soil drying, with evapotranspiration and precipitation data.
6. In another study, it presented the comparison of four regression models for the prediction of soil organic carbon (OC), moisture content (MC), and total nitrogen (TN). They used visible-near infrared (VIS-NIR) spectrophotometer to collect soil spectra.They concluded that the accurate prediction of soil properties can optimize soil management.

Study of Tools and Technology

- **Crop Management**

Crop	Observed Features	Functionality	Models/Algorithms	Results
Wheat	Normalized values of on-line predicted soil parameters and the satellite NDVI	Wheat yield prediction within field variation	ANN/SNKs	81.65% accuracy
Tomato	High spatial resolution RGB images	Detection of tomatoes via RGB images captured by UAV	Clustering/EM	Recall: 0.6066 Precision: 0.9191 F-Measure: 0.7308
Rice	Twenty (20) chemical components that were found in composition of rice samples with inductively coupled plasma mass spectrometry	Prediction and classification of geographical origin of a rice sample	EL/RF	93.83% accuracy

Fig 4.1

- **Water Management**

Property	Observed Features	Functionality	Models/Algorithms	Results
Evapotranspiration	Data such as maximum, minimum, and mean temperature; relative humidity; solar radiation; and wind speed	Estimation of monthly mean reference evapotranspiration arid and semi-arid regions	Regression/MARS	MAE = 0.05 RMSE = 0.07 R = 0.9999
Daily dew point temperature	Weather data such as average air temperature, relative humidity, atmospheric pressure, vapor pressure, and horizontal global solar radiation	Prediction of daily dew point temperature	ANN/ELM	Region case A: MABE = 0.3240 °C RMSE = 0.5662 °C R = 0.9933 Region case B: MABE = 0.5203 °C RMSE = 0.6709 °C R = 0.9877

Fig 4.2

- ## Soil Management

Property	Observed Features	Functionality	Models/Algorithms	Results
Soil drying	Precipitation and potential evapotranspiration data	Evaluation of soil drying for agricultural planning	IBM/KNN and ANN/BP	Both performed with 91–94% accuracy
Soil condition	140 soil samples from top soil layer of an arable field	Prediction of soil OC, MC, and TN	SVM/LS-SVM and Regression/Cubist	OC: RMSEP = 0.062% & RPD = 2.20 (LS-SVM) MC: RMSEP = 0.457% & RPD = 2.24 (LS-SVM) TN: RMSEP = 0.071% & RPD = 1.96 (Cubist)
Soil temperature	Daily weather data: maximum, minimum, and average air temperature; global solar radiation; and atmospheric pressure. Data were collected for the period of 1996–2005 for Bandar Abbas and for the period of 1998–2004 for Kerman	Estimation of soil temperature for six (6) different depths 5, 10, 20, 30, 50, and 100 cm, in two different in climate conditions Iranian regions; Bandar Abbas and Kerman	ANN/SaE-ELM	Bandar Abbas station: MABE = 0.8046 to 1.5338 °C RMSE = 1.0958 to 1.9029 °C R = 0.9084 to 0.9893 Kerman station: MABE = 1.5415 to 2.3422 °C RMSE = 2.0017 to 2.9018 °C R = 0.8736 to 0.9831 depending on the depth
Soil moisture	Dataset of forces acting on a chisel and speed	Estimation of soil moisture	ANN/MLP and RBF	MLP: RMSE = 1.27% R^2 = 0.79 APE = 3.77% RBF: RMSE = 1.30% R^2 = 0.80 APE = 3.75%

Fig 4.3

Summary

- Five ML models were implemented in the approaches on crop management, where the most popular models were ANNs (with most frequent crop at hand—wheat).
- For water management in particular evapotranspiration estimation, two ML models were implemented and the most frequently implemented were ANNs.
- Finally, in the soil management category, four ML models were implemented, with the most popular one again being the ANN model.

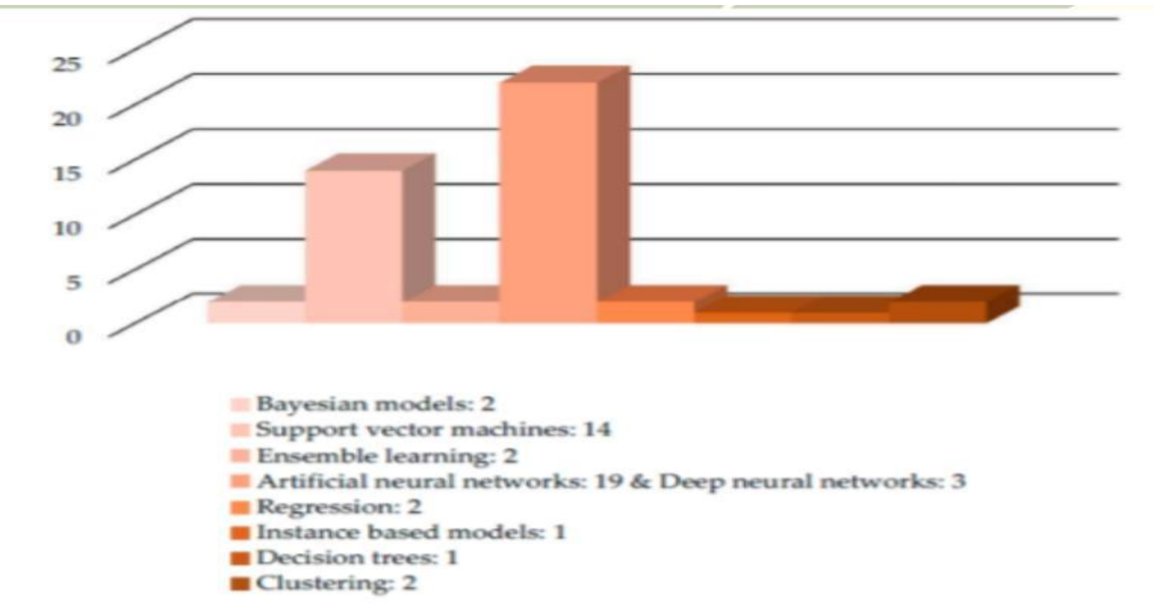


Figure 3. Presentation of machine learning (ML) models with their total rate

Fig 4.4

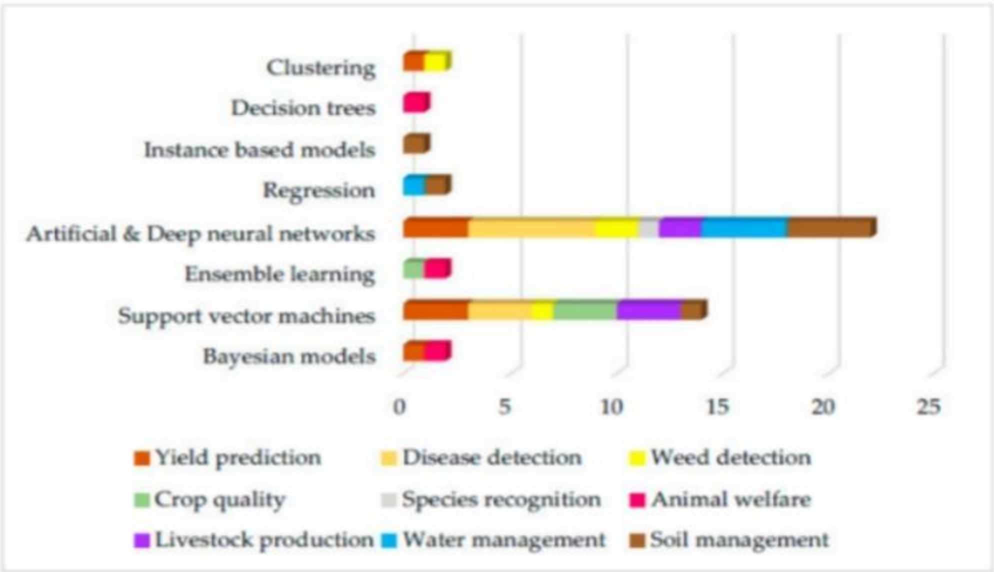


Figure 4. The total number of ML models according to each sub-category of the four main categories.

fig 4.5

V) Yield performance estimation of corn hybrids using machine learning algorithms

Introduction

Estimation of yield performance for crop products is a topic of interest in agriculture. In breeding programs, it is impossible to test hybrids created by crossing two parents (inbred and tester) since it would be too time consuming and costly. In this paper, different machine learning algorithms including decision tree, gradient boosting machine, random forest, adaptive boosting, XGBoost and neural network to predict the yield of corn hybrids using data provided in the 2020 Syngenta Crop Challenge. The participants were asked to predict the yield of missing hybrids which were not tested before. Results show that the prediction obtained by XGBoost is more accurate than other models with a root mean square error equal to 0.0524. Using this approach, identification hybrids with high predicted yield that can be bred to increase corn production.

Related Work

- Recently in 2020,(Elavarasan and Vincent) proposed a recurrent neural network deep learning algorithm over the Q-learning reinforcement learning algorithm to predict yield.. Based on their experiments, this approach can predict yield with an accuracy of 93.7% using environmental, soil, water and crop parameters.
- In another study, a deep convolutional neural network model was developed to extract key features from normalized difference vegetation index (NDVI) and RGB data to predict yield.
- Several empirical breeding programs were developed to explore the impact of the breeding on crop yield and resistance to diseases and environmental factors like temperature and water loss (Duvick et al., 2004).
- In another study researchers identified corn hybrids tolerant to drought and heat using an unsupervised approach based on a stress metrics obtained from a deep learning model (Khaki et al., 2019).

Study of Tools and Technology

- **GBM model-**This variable selection helps to reduce overfitting.In GBM model, we can calculate relative importance for variables which indicates how significant each variable was in the construction of the boosted decision trees within the model.
- Several machine learning models including Adaboost, decision tree, random forest, neural network, XGBoost and GBM were used on the training set and evaluated the models.
- The data were encoded using three methods. The best encoding method was one hot encoding by which we obtained the minimum RMSE among all the models.
- The performance of XGBoost algorithm with integer encoding is slightly better than feature hash encoding. The same pattern holds true for Adaboost and GBM models. Random forest is observed to have better performance compared to Adaboost and decision tree using data encoded by one hot encoding and feature hashing encoding methods.
- The data were encoded using three methods. The best encoding method was one hot encoding by which we obtained the minimum RMSE among all the models.
- The performance of XGBoost algorithm with integer encoding is slightly better than feature hash encoding. The same pattern holds truefor Adaboost and GBM models. Random forest is observed to have better performance compared to Adaboost and decisiontree using data encoded by one hot encoding and feature hashing encoding methods.

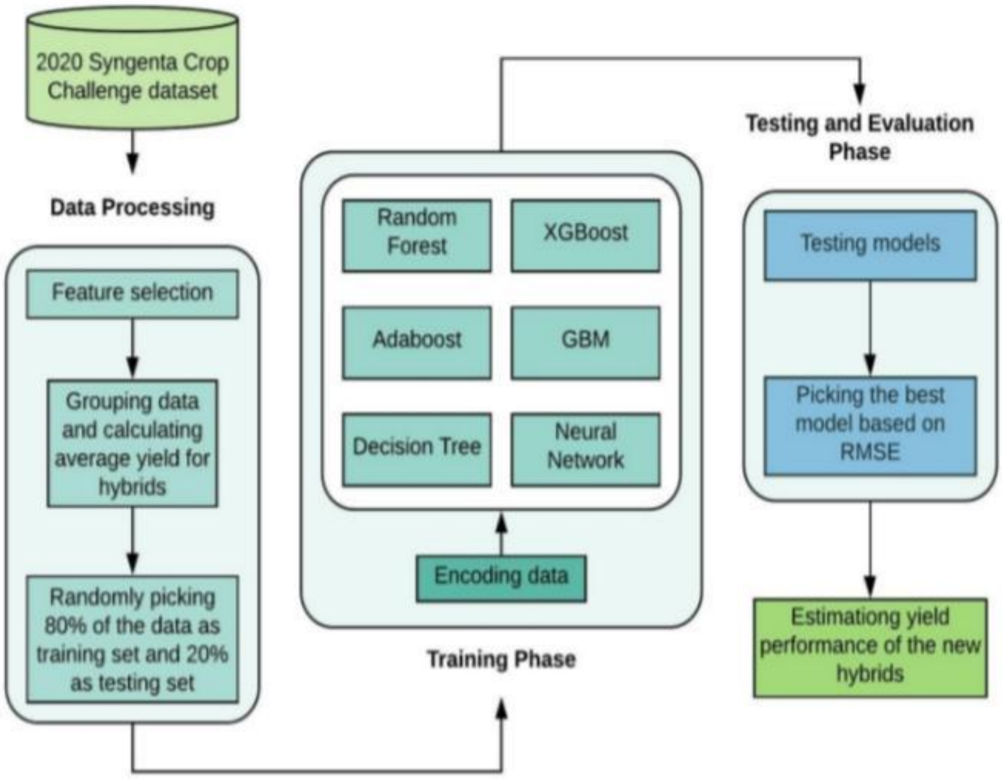


Fig. 2. Flow diagram of the 2020 Syngenta crop challenge analysis; Predicting yield performance of the cross between inbred and tester combinations. This diagram shows the sequence of steps from data processing, model training and testing leading to new hybrids yield prediction.

Fig 5.1

Summary

- This paper presents an application of machine learning algorithms in agricultural data analysis. We implemented various methods on the dataset provided in the 2020 Syngenta crop challenge in order to estimate the performance of new corn hybrids using available tested hybrids. The results in the testing phase implied that XGBoost algorithm has a superior performance among all implemented techniques
- The results demonstrated that more than one third of the new hybrids had indeed a higher yield performance compared to the benchmark in the corresponding environment.

VI)Combining machine learning, space-time cloud restoration and phenology for farm-level wheat yield prediction

Introduction

In this research, researchers created a novel approach by combining four methods. First, they implemented the cloud restoration algorithm called gapfill to restore missed Normalized Difference Vegetation Index (NDVI) values derived from Sentinel-2 sensor (S2) due to cloud obstruction.

Second, they created square tiles as a solution for high computing infrastructure demand due to the use of high-resolution sensor.

Third, they implemented gapfill following critical crop phenology stage.

Fourth, observations from restored images combined with original (from cloud-free images) values and applied for winter wheat prediction. They applied seven base machine learning as well as two groups of super learning ensembles. The study successfully applied gapfill on high-resolution image to get good quality estimates for cloudy pixels. Consequently, yield prediction accuracy increased due to the incorporation of restored values in the regression process.

Related work

- Yield estimation is the process of accurately predicting crop yield before harvest. Remote sensing technologies coupled with empirical methods used to predict crop yield having some comparative advantages, notably, covering a large area, enabling continuous monitoring, and providing timely data (Potgieter et al., 2014; Lobell et al., 2015).
- The relationship between satellite-derived VIs and crop yield is established because important crop characteristics that determine crop yield such as leaf area index, biomass, and chlorophyll content are represented by VIs(reflectance, transmission, and absorbance,biochemical, physiological, and morphological)
- A study revealed that using Quickbird image at end of the heading and in particular after inflorescence fully emerged, NDVI is highly correlated with yield with an average correlation coefficient of 0.77 (Kumhálová and Matějková, 2017)

Study of Tools and Technology

- **Gapfill-**The overall workflow of the gapfill method could be presented in two major parts: the extraction (subset) component and prediction component, which were implemented in sequential order. In R platform, we used packages that include gapfill, raster, foreach, doParallel, and rgdal to support the overall process.
- **Extraction-**This is the first step in which the algorithm iteratively selects a big enough neighborhood which is referred to as a prediction set along the spatio-temporal dimension. The subset is

constructed around the target value (missing value) considered at a time. The row remote sensing data need to be stored in a four-dimensional array and missing values will be stored as NA.

- **Prediction-**Regularized regression (GLMNET) and Quantile regression models were used to as predictive algorithms..
- **Ensemble methods -** Ensemble methods are the third major machine learning approach applied. Ensemble learning is the process of learning from the advantage of combining multiple algorithms for better model performance. In applying a single method of supervised algorithms, the task is to search for a solution in parameter space, while ensembles combine multiple parameter spaces to form a better hypothesis

Summary

- The study evaluated base machine learning and ensemble methods for wheat yield prediction. Ensemble methods produced a similar performance with that of base models. Ensemble models using the seven predictors from base models as well as using three selected predictors based on accuracy following the ranking methodology showed similar performance.
- Overall, satellite-based farm-level yield prediction methods are rarely available in many crop production systems. Therefore, the methods revealed in this study will be crucial in numerous projects such as food security, crop insurance among others where farm-level yield data are key inputs.
- Furthermore, though the study revealed yield prediction methodology for wheat, the same approach could be adopted to develop reliable remote sensing based yield prediction methods for other cereal crops including barley, maize, and sorghum.

VII)Machine Learning: Applications in Indian Agriculture

Introduction

Agriculture is the backbone of every economy. In a country like India, which has ever increasing demand of food due to rising population, advances in agriculture sector are required to meet the needs. To add to it, the present economic conditions and government policies of India are such that it necessitates the adoption of Precision farming or smart farming. It will enable the farmers to maximize their crop yields and minimize the input costs as well as the losses due to reasons like uncertain rainfall, droughts etc. The agriculture sector needs a huge upgradation in order to survive the changing conditions of Indian economy

Related work

- A. Crop Selection and Crop Yield Prediction To maximize the crop yield, selection of the appropriate crop that will be sown plays a vital role. It depends on various factors like the type of soil and its composition,

- A plant nutrient management system has been proposed based on machine learning methods to meet the needs of soil and maintain its fertility levels and hence improve the crop yield . A crop selection method called CSM has been proposed which helps in crop selection based on its yield prediction and other factor
- Disease detection using images of crop leaves has been implemented using pattern recognition branch of machine learning. It works by obtaining patterns from input data and separating them into classes of diseases
- weather prediction based on machine learning technique called Support Vector Machines had been proposed. These algorithms have shown better results over the conventional algorithms and hence have a bright future.

S. No.	Machine Learning Applications in Agriculture		
	Field of Study	**Author**	**Algorithms Used**
1	Crop selection and Crop yield prediction	Washington et al	Classification algorithms
		Snehal et al	Neural Networks
		Shivnath et al	Back propagation neural networks
		Rakesh et al	CSM
2	Weather Forecasting	Y. Radhika et al	Support Vector Machines
3	Smart Irrigation System	Aditya Gupta et al	General Machine learning algorithms
4	Crop disease prediction	Rumpf et al	Support Vector Machines
		M.P. Raj et al	Pattern Recognition
		Mehra et al	Artificial neural networks, Regression trees, Random forests
5	Deciding the MSP	Suggested by author	Classification techniques, Neural networks etc. may be used

Fig 7.1

- A. Directed Learning The calculation is given a few preparation models based on which it can concentrate on the data sources and their relating yields. For instance, showing a youngster the banner of a nation and furthermore letting him know the name of the country it has a place with. On the off chance that the result factors are given, the learning becomes managed. Issues like order and relapse go under this class. Well known administered learning calculations are Artificial neural organizations, Decision trees, K-implies bunching, Support vector machines, Bayesian organizations and so forth
- B. Solo Learning When the calculation isn't furnished with any results, the learning is supposed to be unaided. For instance, assuming we read a book in a language that we don‟t know, we don‟t get anything, yet we continue to peruse or watching we will recognize specific examples in words gradually begin understanding. Calculations including grouping procedures have a place classification. Well known unaided learning calculations Self coordinated component maps, COBWEB, DBSCAN and so forth
- C. Support Learning This sort of learning deals with the standard of input. Each activity affects the framework which is then revealed back to the calculation. The calculation changes its conduct as indicated by the input got. Well known calculations are Genetic calculations, Markov choice calculations and so forth

Summary

Agribusiness is a field that has been inadequate with regards to the mass reception of innovation and its headways. Indian ranchers should be sufficient with the worldwide procedures. AI is an innocent idea that can be very much executed in any field which has complex connections among the info and result factors. It has as of now settled its ability over customary calculations of software engineering and measurements. AI calculations have upgraded the exactness of man-made brainpower machines including sensor based frameworks utilized in accuracy cultivating. This paper has checked on the different uses of AI in the cultivating area. It additionally gives a knowledge into the difficulties looked by Indian ranchers and how they can be settled utilizing these methods.

VIII)Machine learning in agriculture domain: A state-of-art survey

Introduction

Agriculture is considered an important pillar of the world's economy and also satisfies one of the basic need of human being i.e. food. In most of the countries it is considered the major source of employment. Many countries like India still use the traditional way of farming, farmers are reluctant to use advanced

technologies while farming because of either the lack of knowledge, heavy cost or because they are unaware about the advantages of these technologies.
As per the famous saying **"Information is the Power"**, keeping track of information about the crops, environment, and market, may help farmers to take better decisions and alleviate problems related to agriculture.
Technologies like blockchain, IoT, machine learning, deep learning, cloud computing, edge computing can be used to get information and process it. Applications of computer vision, machine learning, IoT will help to raise the production, improves the quality, and ultimately increase the profitability of the farmers and associated domains.

Related Work

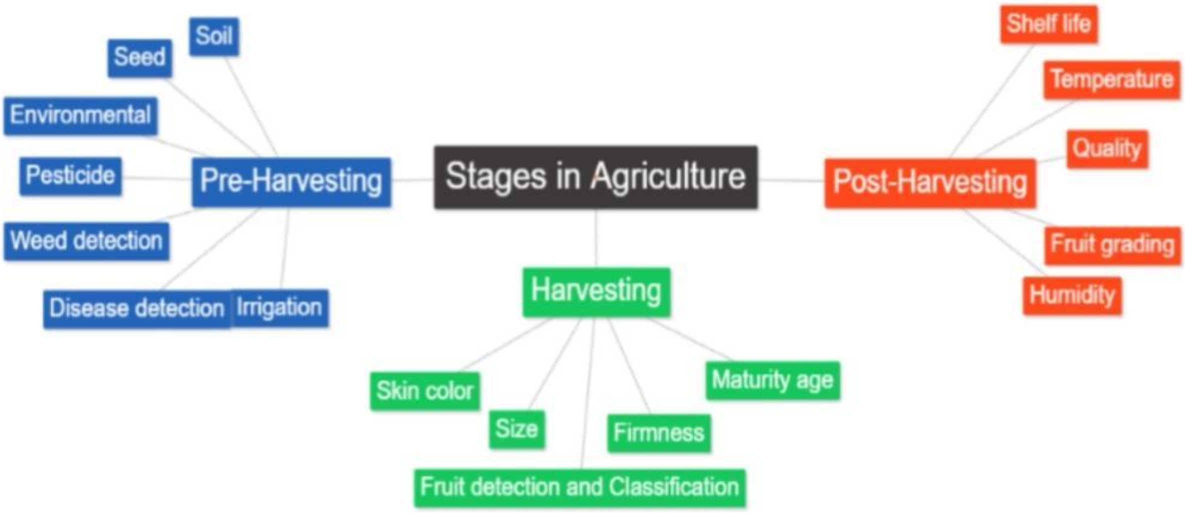


Fig. 2. Important parameters considered in each stage of farming.

Fig 8.1

Study of Tools and Technology

Image preprocessing -which converts image data into a form that allows machine learning algorithms to solve it. It is often used to increase a model's accuracy, as well as reduce its complexity

Feature Extraction-is a part of the dimensionality reduction process, in which, an initial set of the raw data is divided and reduced to more manageable groups. So when you want to process it will be easier. The most important characteristic of these large data sets is that they have a large number of variables.

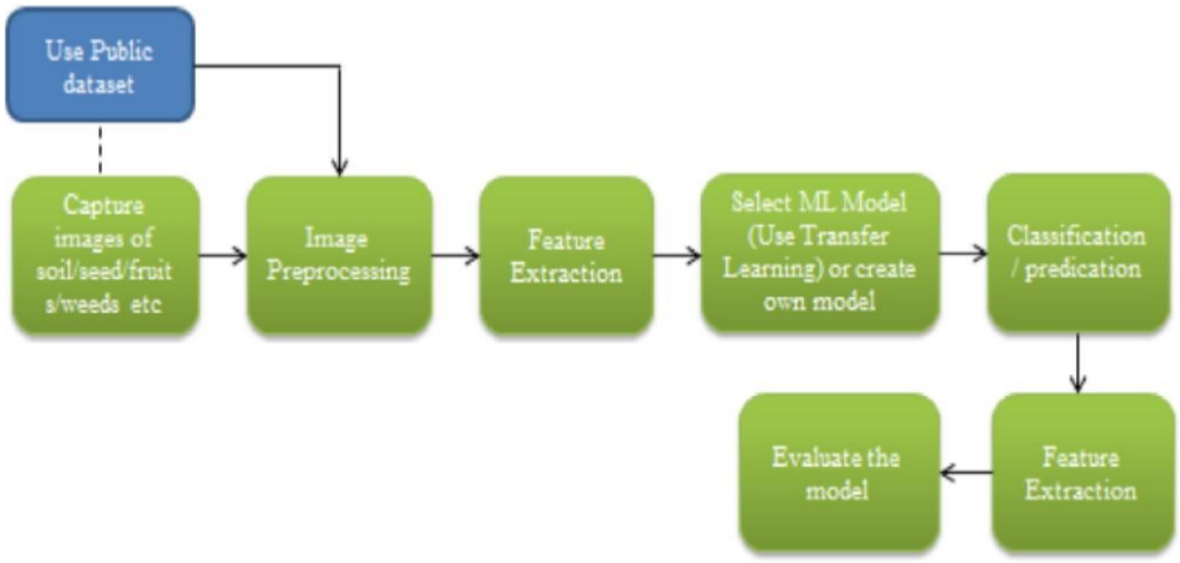


Fig. 3. Steps of Machine Learning used in literature.

Fig 8.2

Summary

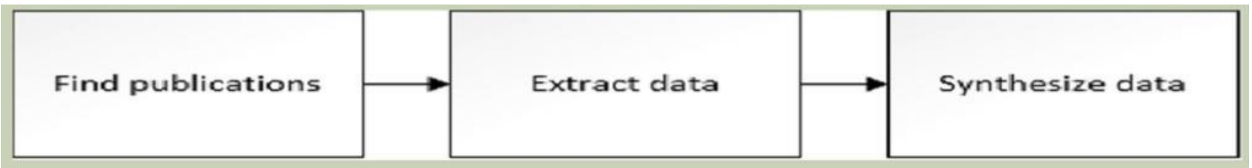
In this paper a top to bottom study of utilizations of AI calculations in agribusiness area is introduced. As per this survey, agribusiness exercises are extensively classified into three significant regions as pre-collecting, gathering and post reaping. Significant boundaries to be considered in each stage are displayed in Fig. 2 and Table 1 . AI calculations/strategies utilized in each stage are evaluated and introduced in Tables 2 , 3 , 4 , 5 and 6 individually. AI is the condition of-workmanship innovation which is utilized to tackle complex issue in the horticulture and assisting ranchers with diminishing their misfortunes. In this review it is seen that AI calculations have acquired surprising results to take care of horticulture related issues.

Introduction

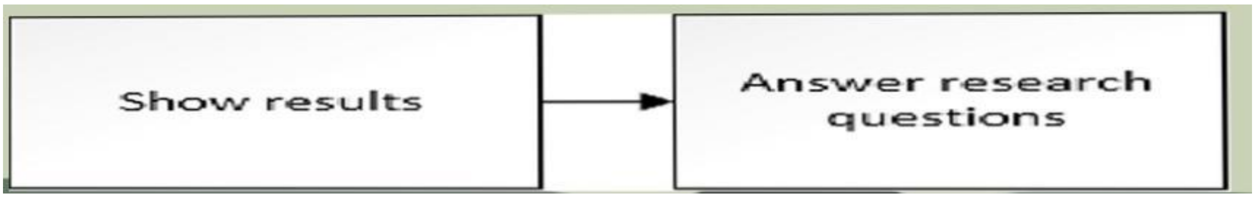
Crop yield prediction is an essential task for the decision-makers at national and regional levels (e.g., the EU level) for rapid decision-making. An accurate crop yield prediction model can help farmers to decide on what to grow and when to grow. There are different approaches to crop yield prediction. This review article has investigated what has been done on the use of machine learning in crop yield prediction in the literature

Related Work

- The first stage is planning the review. In this stage, research questions are identified, a protocol is developed, and eventually, the protocol is validated to see if the approach is feasible.
- The second stage is conducting the review.When conducting the review, the publications were selected by going through all the databases.



- In the final stage, a.k.a., Reporting the Review, the review was concluded by documenting the results and addressing the research questions



Study of Tools and Technology

- As shown in figure the most used features are related to temperature, rainfall, and soil type. Crop yield is the dependent variable.
- To get a better overview of the independent variables the features were grouped.
- The independent features can be grouped into soil and crop information, humidity, nutrients, and field management.
- As shown in this table, the feature groups that are most used are related to the soil, solar, and humidity information.

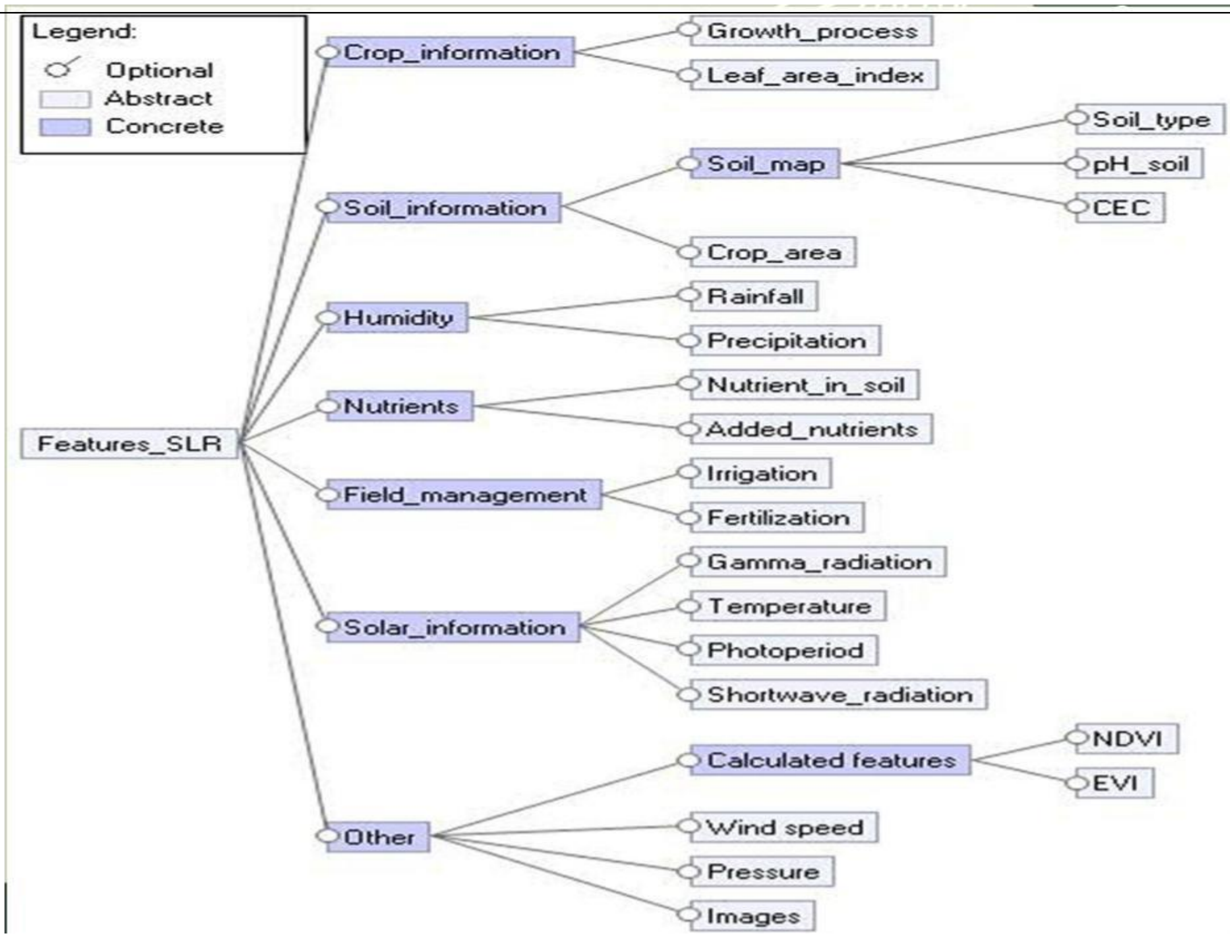


Fig 9.1

Summary

This study showed that the chose distributions utilize an assortment of elements, contingent upon the extent of the exploration and the accessibility of information. Each paper explores yield expectation with AI however varies from the highlights. The examinations additionally contrast in scale, topographical position, and yield. The selection of highlights is subject to the accessibility of the dataset and the point of the exploration. Concentrates on additionally expressed that models with more elements didn't dependably give the best presentation to the yield forecast. To observe the best performing model, models with more and less highlights ought to be tried. Numerous calculations have been utilized in various investigations. The outcomes show that no particular end can be attracted with respect to what the best model is, however they obviously show that some AI models are utilized more than the others. The most utilized models are the arbitrary timberland, neural organizations, direct relapse, and angle helping tree. The majority of the investigations utilized an assortment of AI models to test which model had the best forecast.

X)Implementation of artificial intelligence for optimization of irrigation and application of pesticides and herbicides

Introduction

The advances which are AI-based assistance to further develop productivity in every one of the fields and furthermore deal with the difficulties looked by different businesses remembering the different fields for the rural area like the harvest yield, water system, soil content detecting, crop-checking, weeding, crop foundation. Farming robots are implicit request to convey high esteemed utilization of AI in the referenced area.. Man-made intelligence based mechanical arrangements has empowered the ranchers to deliver more result with less information and surprisingly worked on the nature of result, likewise guaranteeing quicker go-to-advertise for the yielded crops. By 2020, ranchers will utilize 75 million associated gadgets.

Relevant Work

Image recognition and perception

An increasing interest has been seen in autonomous UAVs and their applications including recognition and surveillance, human body detection and geolocalization, search and rescue, forest fire detection.

Skills and workforce

Artificial intelligence makes it possible for farmers to assemble large amount of data from government as well as public websites, analyze all of it and provide farmers with solutions to many ambiguous issues as well as it provides us with a smarter way of irrigation which results in higher yield to the farmers.

Chatbots for farmers

Chatbots are nothing but the conversational virtual assistants who automate interactions with end users. Artificial intelligence powered chatbots, along with machine learning techniques has enabled us to understand natural language and interact with users in away more personalized way.

Study of Tools and Technology

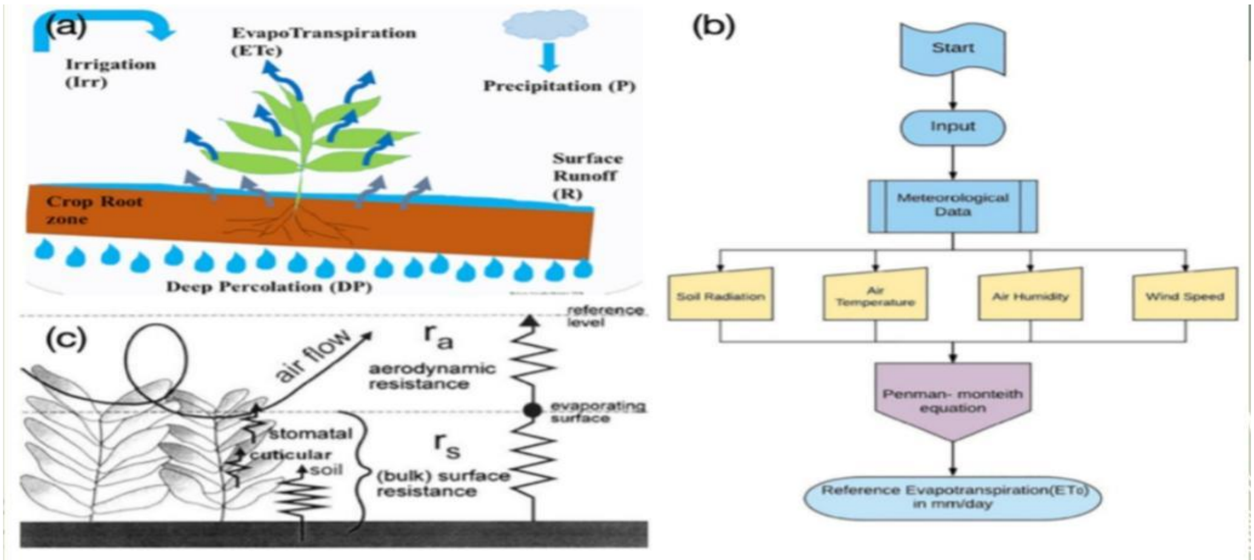


Fig 10.1

Summary

The agrarian business faces different difficulties, for example, absence of viable water system frameworks, weeds, issues with plant checking because of yield stature and outrageous climate conditions. However, the presentation can be expanded with the guide of innovation and hence these issues can be settled. It very well may be improved with various AI driven procedures like far off sensors for soil dampness content identification and computerized water system with the assistance of GPS. The issue looked by ranchers was that accuracy weeding strategies beat the enormous measure of harvests being lost during the weeding system. Not exclusively do these independent robots further develop productivity, they additionally decrease the requirement for superfluous pesticides and herbicides. Other than this, ranchers can shower pesticides and herbicides viably in their homesteads with the guide of robots, and plant observing is additionally presently not a weight. First of all, deficiencies of assets and occupations can be perceived with the guide of man-made mental ability in agribusiness issues. In customary techniques tremendous measure of work was needed for getting crop qualities like plant stature, soil surface and content, as such manual testing happened which was monotonous.

CONCLUSION

India is a country where farming assumes a great part. In success of the ranchers, flourishes the country. Hence our work would help ranchers in planting the right seed in light of soil prerequisites to build usefulness of the country. Our future work is focused on a better informational collection with enormous number of traits and furthermore carries out yield forecast.

References:-

[1]. SatishBabu (2013),'A Software Model for Precision Agriculture for small and Marginal Farmers'. At the International Centre for and Open Source Software(ICFOSS) Trivandrum, India.

[2]. AnshalSavla, ParulDhawan, HimtanayaBhadada, NiveditaIsrani, Alisha Mandholia, SanyaBhardwaj (2015),'Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture', Innovation in Information, Embeddedand Communication system(ICIIECS).

[3]. Aakunuri Manjula, Dr. G. Narsimha (2015),'XCYPF:AFlexibleand Extensible Framework for Agriculture Crop Yield Prediction', Conference on Intelligent Systems and Control(ISCO).

[4] Rakesh Kumar , M.P. Singh, Prabhat Kumar and J.P. Singh, “Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique”, International Conference on Smart Technologies and Management for Computing, Communication, Controls,Energy and Materials, 2015

[5] Haedong Lee and Aekyung Moon, “Development of Yield Prediction System Based on Real-time Agricultural Meteorological Information”, 16th International Conference on Advanced Communication Technology, 2014

[6] T.R. Lekhaa, “Efficient Crop Yield and Pesticide Prediction for Improving Agricultural Economy using Data Mining Techniques”, International Journal of Modern Trends in Engineering and Science(IJMTES), 2016, Volume 03, Issue 10

[7] Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade and Vahida Attar, “Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction”, International Journal of Computer Science Issues, Volume 9, Issue 3

[8] Anshal Savla, Parul Dhawan, Himtanaya Bhadada, Nivedita Israni, Alisha Mandholia, Sanya Bhardwaj (2015), ‘Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture’, Innovations in Information, Embedded and Communication systems (ICIIECS).

[9] AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms. Zeel Doshi , Subhash Nadkarni , Rashi Agrawal, Prof. Neepa Shah.

[10] Crop Recommendation System for Precision Agriculture
S.Pudumalar*, E.Ramanujam*, R.Harine Rajashree, C.Kavya,

T.Kiruthika, J.Nisha.

[11] Tom M. Mitchell, Machine Learning, India Edition 2013, McGraw Hill Educa- tion.

[12] https://data.gov.in/g

[13] Kagglehttps://www.kaggle.com/notebook
14. Bohanec, M.; Kljaji´c Borštnar, M.; Robnik-Šikonja, M. Explaining machine learning models in sales predictions.
Expert Syst. Appl. 2017, 71, 416–428. [CrossRef]
15. Takahashi, K.; Kim, K.; Ogata, T.; Sugano, S. Tool-body assimilation model considering grasping motion
through deep learning. Rob. Auton. Syst. 2017, 91, 115–127. [CrossRef]
16. Gastaldo, P.; Pinna, L.; Seminara, L.; Valle, M.; Zunino, R. A tensor-based approach to touch modality classification by using machine learning. Rob. Auton. Syst. 2015, 63, 268–278. [CrossRef]
17. López-Cortés, X.A.; Nachtigall, F.M.; Olate, V.R.; Araya, M.; Oyanedel, S.; Diaz, V.; Jakob, E.; Ríos-Momberg, M.; Santos, L.S. Fast detection of pathogens in salmon farming industry. Aquaculture 2017, 470, 17–24. [CrossRef]
18. Zhou, C.; Lin, K.; Xu, D.; Chen, L.; Guo, Q.; Sun, C.; Yang, X. Near infrared computer vision and neuro-fuzzy
model-based feeding decision system for fish in aquaculture. Comput. Electron. Agric. 2018, 146, 114–124. [CrossRef]
19. Fragni, R.; Trifirò, A.; Nucci, A.; Seno, A.; Allodi, A.; Di Rocco, M. Italian tomato-based products authentication by multi-element approach: A mineral elements database to distinguish the domestic provenance. Food Control 2018, 93, 211–218. [CrossRef]
20. Maione, C.; Barbosa, R.M. Recent applications of multivariate data analysis methods in the authentication of
rice and the most analyzed parameters: A review. Crit. Rev. Food Sci. Nutr. 2018, 1–12. [CrossRef] [PubMed]
21. Fang, K.; Shen, C.; Kifer, D.; Yang, X. Prolongation of SMAP to Spatiotemporally Seamless Coverage of
Continental U.S. Using a Deep Learning Neural Network. Geophys. Res. Lett. 2017, 44, 11030–11039. [CrossRef]
22. Pearson, K. On lines and planes of closest fit to systems of points in space. Lond. Edinb. Dublin Philos. Mag.
J. Sci. 1901, 2, 559–572. [CrossRef]
23. Wold, H. Partial Least Squares. In Encyclopedia of Statistical Sciences; John Wiley & Sons: Chichester, NY, USA,
1985; Volume 6, pp. 581–591, ISBN 9788578110796.
24. Fisher, R.A. The use of multiple measures in taxonomic problems. Ann. Eugen. 1936, 7, 179–188. [CrossRef]
25. Cox, D.R. The Regression Analysis of Binary Sequences. J. R. Stat. Soc. Ser. B 1958, 20, 215–242. [CrossRef]
26. Efroymson, M.A. Multiple regression analysis. Math. Methods Digit. Comput. 1960, 1, 191–203. [CrossRef]
27. Craven, B.D.; Islam, S.M.N. Ordinary least-squares regression. SAGE Dict. Quant. Manag. Res. 2011, 224–228.
28. Friedman, J.H. Multivariate Adaptive Regression Splines. Ann. Stat. 1991, 19, 1–67. [CrossRef]

29. Quinlan, J.R. Learning with continuous classes. Mach. Learn. 1992, 92, 343–348.
30. Cleveland, W.S. Robust locally weighted regression and smoothing scatterplots. J. Am. Stat. Assoc. 1979, 74,
829–836. [CrossRef]
31. Tryon, R.C. Communality of a variable: Formulation by cluster analysis. Psychometrika 1957, 22, 241–260.
[CrossRef]
32. Lloyd, S.P. Least Squares Quantization in PCM. IEEE Trans. Inf. Theory 1982, 28, 129–137. [CrossRef]
33. Johnson, S.C. Hierarchical clustering schemes. Psychometrika 1967, 32, 241–254. [CrossRef] [PubMed]

34. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm.
J. R. Stat. Soc. Ser. B Methodol. 1977, 39, 1–38. [CrossRef]
35. Russell, S.J.; Norvig, P. Artificial Intelligence: A Modern Approach; Prentice Hall: Upper Saddle River, NJ, USA,
1995; Volume 9, ISBN 9780131038059.
36. Pearl, J. Probabilistic Reasoning in Intelligent Systems. Morgan Kauffmann San Mateo 1988, 88, 552.
37. Duda, R.O.; Hart, P.E. Pattern Classification and Scene Analysis; Wiley: Hoboken, NJ, USA, 1973; Volume 7,
ISBN 0471223611.
38. Neapolitan, R.E. Models for reasoning under uncertainty. Appl. Artif. Intell. 1987, 1, 337–366. [CrossRef]
39. Fix, E.; Hodges, J.L. Discriminatory Analysis–Nonparametric discrimination consistency properties.
Int. Stat. Rev. 1951, 57, 238–247. [CrossRef]
40. Atkeson, C.G.; Moorey, A.W.; Schaalz, S.; Moore, A.W.; Schaal, S. Locally Weighted Learning. Artif. Intell.
1997, 11, 11–73. [CrossRef]
41. Kohonen, T. Learning vector quantization. Neural Netw. 1988, 1, 303. [CrossRef]
42. Belson, W.A. Matching and Prediction on the Principle of Biological Classification. Appl. Stat. 1959, 8, 65–75.
[CrossRef]
43. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. Classification and Regression Trees; Routledge: Abingdon,
UK, 1984; Volume 19, ISBN 0412048418.
44. Kass, G.V. An Exploratory Technique for Investigating Large Quantities of Categorical Data. Appl. Stat. 1980,
29, 119. [CrossRef]
45. Quinlan, J.R. C4.5: Programs for Machine Learning; Morgan Kaufmann Publishers Inc.: San Francisco, CA,
USA, 1992; Volume 1, ISBN 1558602380.
46. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys.
1943, 5, 115–133. [CrossRef]
47. Broomhead, D.S.; Lowe, D. Multivariable Functional Interpolation and Adaptive Networks. Complex Syst.
1988, 2, 321–355.
48. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain.

Psychol. Rev. 1958, 65, 386–408. [CrossRef] [PubMed]
49. Linnainmaa, S. Taylor expansion of the accumulated rounding error. BIT 1976, 16, 146–160. [CrossRef]
50. Riedmiller, M.; Braun, H. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In Proceedings of the IEEE International Conference on Neural Networks, San Francisco, CA, USA, 28 March–1 April 1993; pp. 586–591. [CrossRef]
51. Hecht-Nielsen, R. Counterpropagation networks. Appl. Opt. 1987, 26, 4979–4983. [CrossRef] [PubMed]
52. Jang, J.S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Trans. Syst. Man Cybern. 1993,
23, 665–685. [CrossRef]
53. Melssen, W.; Wehrens, R.; Buydens, L. Supervised Kohonen networks for classification problems. Chemom. Intell. Lab. Syst. 2006, 83, 99–113. [CrossRef]
54. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA 1982, 79, 2554–2558. [CrossRef] [PubMed]
55. Pal, S.K.; Mitra, S. Multilayer Perceptron, Fuzzy Sets, and Classification. IEEE Trans. Neural Netw. 1992, 3,
683–697. [CrossRef] [PubMed]
56. Kohonen, T. The Self-Organizing Map. Proc. IEEE 1990, 78, 1464–1480. [CrossRef]
57. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme learning machine: Theory and applications. Neurocomputing
2006, 70, 489–501. [CrossRef]
58. Specht, D.F. A general regression neural network. IEEE Trans. Neural Netw. 1991, 2, 568–576. [CrossRef]
[PubMed]