Journal of Engineering, Science and Mathematics

Volume 02, Number 01, Pages 40-56, Year 2023
https://jesm.in/archives/

Research Paper

Ensuring Food Security and Agriculture Demand using Machine Learning and Sensor based Technologies

Rijwan Khan[1], Santosh Kumar[2*], Ankur Seem[3], Arpit Kumar Chauhan[4], Anubhav Gupta[5]

[1,3,4,5]ABES Institute of Technology, Ghaziabad, Affiliated to AKTU Lucknow, India

[2]University of Dar Es Saalam, Tanzania

Correspondence should be addressed to Santosh Kumar; drsengar2002@gmail.com

Handling Editor: Evans Asenso

Agriculture is important for sustaining human life. Contribution of agriculture to the Indian economy is around 20%. It is estimated that by the year 2050, Indian will need 60% more food to feed the population of 9.3 billion. However, due to limited resources and land, there is a challenge of food insecurity, need to enhance the efficiency of current farms, plan, smartly grow crops according to the demand to decrease wastage and ensure food security. Technologies like Machine Learning, IoT, Blockchain, Data Analytics, Big Data, and Cloud Computing hold the key to solving this problem. With these technologies, the authors can analyze the real-time and past data of agriculture and make the best decision for problems like crop selection, demand prediction, weather prediction, and many more. Machine Learning algorithms use data and by doing complex computation, try to give an accurate result. In this paper, authors review different research done in the field of food security agriculture using ML (Machine Learning) by using past data and avoiding live data, which makes the model more affordable by decreasing the cost of IoT devices needed for live data.

Keywords: *ML (Machine Learning), Deep Learning, SVM (Support Vector Machine), Food Security, CNN (Convolutional Neural Network)*

1. Introduction

Due to the increasing population, the demand for food is growing every day and due to the same reason, the land for agriculture is constantly decreasing, which poses a major threat to food security. Large amount of food which is produced

in the present time can get wasted due to many factors. Need to pay attention not only to the production of food but also the proper management and distribution of the food produced. Therefore, we need to adopt methods which will help in increasing the efficiency of agricultural land and help in managing the produced food properly. This approach is named as precision agriculture [1]. Modern agriculture practices enhance the quantity of food, but they also pose the threat to the environment and sometimes also decrease the quality of food which may harm the consumer. There is a need to bring sustainable agriculture practices into effect. With the help of machine learning, it is very much possible to bring the theoretical practices in real life. ML (Machine Learning) provides a way to compute the already present data of agriculture or related fields and give the result which can be used for adopting better practices. In ML, image processing is a large blessing which is used for identifying diseases and distinguishing crops from weeds which can save a large amount of time and effort of farmers. Data prediction is also an important part of ML which helps in other areas like crop yield prediction, crop selection. Customer analysis and demand prediction are also essential for ensuring food security and preventing from food insecurity [2]. The current method of agriculture and supply chain ecosystem will not be able to meet the future demand of population and if the current situation continues, it may lead to the lack of interest of farmers in farming due to the exploitation of farmers. To prevent this situation, the nation needs to create a smart ecosystem with the help of ML which will be able to increase the production, manage the produced crop properly, meet the demand properly, and give the best price to the farmer to motivate them towards farming [3]. In this research, a supply chain model is introduced for the farmers to increase their income. Small Industries and buyers can directly approach to farmers for their crops.

2. Literature Survey

2.1 Machine learning in Crop Selection and Yield Prediction

At present time, there are many crops in India which are produced in large quantity which even exceeds the demand of population by the large amount due to which a portion of that crop gets wasted. Whereas there are many crops which import by paying a large amount of money, but these crops can be easily grown in our country. There is a large gap between what farmers grow and what they need [4]. Our farmers need to be guided about what to grow. Many people have done research on how authors can use different ML algorithms to help farmers in selecting the crop. The Linear Programming based optimization was shown for recommending the most excellent crop on the basis of internal demand, export scope, and many other constraints like availability of seeds, nutrition, labor, water storage, etc. The author presented two case studies. Firstly, dividing a single farmland into 8 fields and using 3 crops. The second case uses four different farmlands in different climate zones which make it more complex to analyze and suggest crops. In the paper, the authors focused on how to maximize the profit of farmers by suggesting the best crop [5]. The fluffy multicriteria decision-making strategy MULTIMOORA is used for phonetic and numeric prioritization of vitality crops beneath Lithuanian climatic conditions. The research was done for finding the alternative energy sources of fossil fuels in the form of energy crops. In

the research, the first crops were ranked on the basis of a fuzzy ratio system. Secondly, the crops were ranked on the basis of their deviation from the fuzzy Maximal Objective Reference Point (MORP). The last ranking is done on the basis of the fuzzy Multiplicative Form [6]. Multicriteria decision analysis (MCDA) permitted the evolution or modification of diverse crop species as a cover crop in Spain according to their suitability. The variables considered were crop biomass, N derived, ground cover, N uptake, C/N, crop residual quality, and dietary fiber content. Use of a MCDA weighted normal approach permits positioning the execution of different species and cultivars, concurring their capacity to achieve a few of the most cover crop capacities [7]. An algorithm called crop selection method (CSM) everywhere and assisting in the choosing of crops depending on various factors is proposed. Authors consider rainfall as the key factor for upcoming year agriculture. They gave large importance to rainfall and previous year data. Depending on this information, this algorithm suggests the best combination of crops with high crop yield for the whole year [8]. A unique way for the selection of crops in Karnataka is proposed, as they consider only five factors, i.e., rainfall, soil moisture, humidity, temperature, and slope. In the front-end user enters the location, the corresponding nearest longitude and latitude of the location are used for further calculation. For each crop, the preference of these five factors was marked allotting each crop code depending on the preference of the factors. Two lists were formed, the main list contains all crops having the top two preferences fulfilled for a particular coordination, and the second list contains all crops whose third preference was satisfied. If the first list turns out to be empty, then the second list crops were shown as a result. If the second list comes out to be empty, then it was considered that more data on different crops are needed for the particular area [9]. The yield prediction of crops using supervised is used in this research. They considered rainfall, ph, humidity, temperature, and crop name as the main factors for yields for a crop. According to them, yield prediction plays important for a farmer in predicting the cost for the farmer. With the help of crop yield prediction, a reduction in loss could be possible. The dataset used for the research was taken from Kaggle. Dataset comprises of variables like ph, humidity, temperature, and rainfall. During the research, it was found out that random forest was more suitable in the comparison decision tree and gives more accurate results [10]. In this research, considering soil as the only parameter, they focused on soil crop, soil characteristics, and crop yield. Datasets used in this research had characteristics of soil like ph, texture, water-storing, soil texture, permeability, drainage, depth, and erosion. In the paper, they explained how these characteristics affect the growth of the crop, what kind of crops were possible to grow on such soil, and which crop gives the maximum profit. For crop suggestion, algorithms like Random tree, CHAID, K-Nearest Neighbor, and Naïve Bayes were used and the results of all algorithms were used and with the help of the majority voting technique, the best crop was selected [11]. In this paper, three algorithms, Artificial Neural Network (ANN), Support Vector Machine (SVM), Decision Forest, were used for different crops. The crops selected for the experiment were rice, potatoes, and wheat. The parameters considered for the experiment were humidity, urea, soil, rainfall, and temperature. For rice, accuracy of ANN, SVM, and RF (Random

Forest) accuracy were 96.4, 73.3, and RF 90.7. Similarly, for potatoes, the accuracies from ANN, SVM, and RF were 96.1, 65.2, and 88.7 and for wheat the respective accuracies from ANN, SVM, and RF were 96.4, 66.3, and 90. Therefore, from the results, it is clear that ANN was best suited for crop prediction and SVM was the least favourable [12]. Weather conditions and soil health highly affect the crop and a recommended ANN for crop prediction is proposed. They consider PH, nitrogen, temperature, and rainfall as the main factors and suggested ANN as a powerful tool for modelling and prediction. The suggested model uses feed-forward back-propagation, feed-forward is used to recall patterns and processes, and back-propagation is used to decrease the error in the model [13]. In this paper, the authors compared two data mining techniques, multiple regression and Density-Based spatial clustering. In Multiple Linear Regression (MLR), multiple independent parameters are utilized to urge the result of the dependent variable. In this research, the main independent characteristics considered were year, rainfall, nitrogen, potassium, phosphorous, area of sowing, and yield. Using all these independents, the future yield of the crop is dependent, which is calculated with the help of MLR. In Density-Based Clustering, the neighborhood of a given point should have at least a minimum number of points within a unit distance. Results generated from these both algorithms were compared [14]. In Table 1, we found some ML models for crop selection and yield prediction. Table has two columns, in the first column ML algorithms are written and in the second column references.

Table 1. Table 1 ML models for Crop Selection and Yield Prediction

ML Algorithm	Research Topic
[15] Linear Programming based optimization model	A Decision Support Framework for National Crop Production Planning.
[6] MULTIMOORA	Fuzzy decision support methodology for sustainable energy crop selection
[16] Multi-criteria Decision Analysis (MCDA)	Multi-criteria decision analysis was applied to cover crop species and cultivar selection.
[17] Supervised learning	Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector
[11] CHAID	Crop recommendation system for precision agriculture.
[18] ANN, SVM, Decision Forest	ML Methods for Crop Yield Prediction and Estimation: An Exploration
[19] ANN	Agricultural crop yield prediction using artificial neural network approach
[20] DBC (Density-Based Clustering) and MLR (Multi Linear Regression)	Analysis of crop yield prediction using data mining techniques

2.2. Machine Learning in Weather Prediction

During the agricultural production phase, weather forecasting plays an essential role. Future farming will require weather forecasting, as well as sophisticated and time-based data collection. The suitable utilization of water for crop water system planning and arranging is guided by climate estimates such as daylight, precipitation, stickiness, and dampness levels. For weather forecasting, machine learning algorithms like supervised and unsupervised methods can be used [21]. As Compared two algorithms for weather prediction algorithm. The first algorithm

was PSO (Particle Swarm Optimization) with Multi-Layer Perceptron (MLP), the purpose of this algorithm was not only to use for rain forecasting but to also enhance the network. Another algorithm used was Levenberg Marquardt with backpropagation. Both these algorithms were compared on the basis of RMSE. RSME (Root Mean Square Error) for PSO was 0.139 and for LM was 0.181. From this we can conclude that in this case, PSO performed better than LM (Levenberg Marquardt) [22]. In this research, the predict of weather is explained and then tried to use that data for crop selection. In this paper, RNN (Recurrent Neural Network) was applied for weather prediction and RF technique was applied for the choosing of good crops. For weather forecast, factors like temperature, humidity, wind speed, wind direction, sun hours, etc were considered. RNN uses all this data with LSTM (Long Short-Term Memory), which helps the RNN model in retaining previously calculated information to be utilized in later steps. Weather data predicted from the RNN model along with soil properties was used in Random Forest and a suggestion of crop was made [23]. Weather radar echo image is used as a main reference for weather forecast, it is exceptionally critical for the soundness of rural generation. Disastrous weather is represented by different radar echo patterns such as harsh convection, hail, and further on. Climate radar echo shape forecast can assist weather forecasters in predicting coming time changes in hazardous climates, avoiding the damage that harsh climate can cause to agribusiness, and reducing agricultural financial losses [24]. Here, Table 2 represents ML models in some recent research on weather prediction.

2.3 Machine Learning in Price and Demand Prediction

The time sequence and Auto-Regressive Integrated Moving Average (ARIMA) method were used for price forecast of crops. Factors considered for the prediction system were food demand, feed demand, climate, soil condition, the current situation of market price, and yearly yield prediction. The algorithm shows the potential to predict at most 95% confidence in its prediction [25]. Nine algorithms for energy demand prediction were compared. Some of the algorithms were SVM, Linear Regression, Gaussian Process, Regression tree, Multi-Layer Perceptron, and many more. It was concluded that deep learning enhanced MLP gives better prediction accuracy.

Table 2. ML model for Weather Prediction

ML algorithms	**Research Manuscripts**
[21] Supervised and Unsupervised	Big data in precision agriculture through ICT: rainfall prediction using neural network approach
[22] PSO with MLP (Multi-Layer Perceptron)	Hybrid Machine Learning Model for Rainfall Forecasting
[23] Recurrent Neural Network (RNN), Random Forest	Machine Learning convergence for weather-based crop selection

It gives the best accuracy in comparison with other eight algorithms in terms of Probability of Correct Classification (PCC), Mean Absolute Percentage Error (MAPE), Normalized Root Mean Square Error (NRMSE) and Root Mean Square Error (RMSE) [26]. Different algorithms like neural network, linear regression, XF Boost, decision tree, and clustering are used for price prediction of soybean in China. Authors considered growth area, yield/production, and export/import as important factors for price prediction of soybean for the local market. The authors

used neural linear regression for finding relevant data from the dataset. Neural Network (NN) is used for price or fare predicting to enhance the precision and RMSE (root mean square error) has been used for measuring the accuracy of each model. From the calculation, it was concluded that XG Boost predicts the target values better in comparison to any other algorithm [27]. A smart short-term price prediction (approx. for 3 months) model was developed. The authors compared two models MA (Moving Average) and ARIMA (Autoregressive Integrated Moving Average). To calculate the accuracy MAPE (Mean Absolute Percentage Error) and RMSE (Root Mean Square Error) were used. Prediction has been applied to different products, among which DDP30 and AOTG34LFT gave the best result. From the results, it was concluded that ARIMA and MA, ARIMA gave the best result for interim prediction, eventually this model is not appropriate for long-term prediction [28]. Here, Table 3 represents ML models in some recent research on demand prediction.

Table 3. ML model for Price and Demand Prediction

ML algorithms	Research
[25] ARIMA (Auto-Regressive Integrated Moving Average)	Price Forecasting System for Crops at the Time of Sowing
[26] SVM, Linear Regression, Gaussian Process, Regression tree, Multi-Layer Perceptron	Deep learning versus traditional machine learning methods for aggregated energy demand prediction
[27] Neural network, linear regression, XF Boost, decision tree, and clustering	Crop Price Prediction System using Machine learning Algorithms
[28] Moving Average and Autoregressive Integrated Moving Average	A smart system for short-term price prediction using time series models

2.4 Machine Learning and IoT Sensors in Food security

Food security includes always having the material and financial resources to have enough food to encounter the dietetic require for a healthy and beneficial life. The family is guaranteed food when its members are not hungry or afraid of starving to death. Food fragilities are often linked to deprivation and have lasting effects on the ability of families, communities, and nations to create and thrive. [29]. Growth retardation retards growth, moderately improves cognition, and increases helplessness in the face of disease. Today, more than 800 million people around the world go to be hungry every night, most of them small-scale herders who rely on agribusiness to make ends meet and support their families [30]. Despite the explosion of urban slum development over the past decade, about 75% of the poor in developing countries live in rural areas. Development in the horticultural sector, from crop to fork, appears to be at least twice as effective in reducing poverty as development in other areas. Contributing to those smallholder farmers, many of whom are women and the nourishment systems which feed them are more vital than ever. Authors arrange for bolster a populace anticipated to develop with 9 billion individuals by 2050, the world ought to two fold its current nourishment generation. Given the scarcity of characteristic assets and other challenges, the world will get to be more productive in how it meets this request. To guarantee that individuals have adequate nourishment, adjusting short-term help with a long-term advancement procedure

can offer assistance between nations to nourish their possessed individuals [31]. In the paper [32], the authors created a conceptual specialized guide of autonomous pollination for future cultivating utilizing automated miniaturized scale air vehicle pollinators. The research gives modern bits of knowledge into an independent plan and makes into conceivable different ideas to extend generation effectiveness, abbreviated time from lab to showcase. An independent miniaturized scale air vehicle pollinators are realized utilizing manufactured intelligence and human ability within the circle of the shrewd agrarian industry. Encouraged, this work distinguishes logical and innovative progresses that are anticipated to decipher, inside the proposed administrative system, into the inescapable utilize of miniaturized scale air vehicle pollinators for agrarian applications and past. In this research, independent fertilization is one of the basic advances for exactness cultivated within the close future. For future exactness cultivating, fertilization will carry out one of its central exercises and mindful to 90% creatures that live on our planet, to what end it may be an exchange of dust from the anther to the stigma. Not many plants might be pollinated either way by air or creatures, but nearly all plants may be not competent for independent self-pollination. When pollination is required on a large scale in future sustainable cultivating, for illustration, field crops, plantations, or commercial seed generation, unused advances must be investigated to overcome this challenge on worldwide future nourishment security. A robotic micro air vehicle (MAV) pollinator (MPr) is one of the rising arrangements to support worlds' food supply. This paper points to supply a conceptual specialized guide of an independent fertilization framework for future cultivating utilizing robotic micro discuss vehicle pollinators with fluttering wings. Micro air vehicles with fluttering wings have been altogether created amid the later long time, owing to their characteristics of little size, vitality, proficiency, and dexterity. They offer a wide run of potential civilian and military applications. Therefore, distant, there are a few fruitful fluttering wing Micro air vehicles, such as Nano Humming-bird [32] by Aerovironment Inc., DelFly [33] by TU Delft, Harvard Microrobotic Fly [34], Harvard RoboBees, Robot Dragonfly by TechJect, FESTO BionicOpter and Mosquito robot, etc. Most of the research work has centered on robot basic plan, circuit plan, or flight recreations, the less (FESTO) have tended to the 'intelligence' of the 'autonomous' automated frameworks. To fill this crevice, this paper proposes an independent automated pollinator utilizing a computational insight approach. The commitment of this work can be laid out as a specialized guide of independent pollination for future cultivating employing an automated micro air vehicle framework with AI-in-the-loop (Trouble) and Human-in-the-loop (HIL). Each micro air vehicles of the mechanical swarm are controlled by the central control framework (CCS) by means of remote flag associations. Real time field information, such as trim and blooms, are captured by cameras and other sensors (e.g., thermal sensors) in the central control framework. It cleverly central control framework creates the control signals. The framework is outlined through computational insights (CI), which could be a set of nature-inspired approaches advertising a rich range of capabilities for complex problem solving. Compared to the traditional optimization strategies, computational insight does not have to be reformulated for the issue to look a nonlinear or non-

differentiable space. To arrange to use the benefits of a differing extent of CI approaches, the computational insight coordinates solver has been used in this inquire about [35], [36].

In a research paper [37], the authors used two diverse countries-sharpened French and Moroccan honeys from diverse topographical regions were classified and characterized by applying a voltammetric electronic tongue (VE-tongue) coupled to expository strategies. The considered parameters incorporate color concentrated, free lactonic and add up to corrosiveness, proteins, phenols, hydroxymethyl furfural substance (HMF), sucrose, diminishing and add up to sugars. The topographical classification of distinctive honeys was created through three pattern acknowledgment procedures: central component examination (PCA), bolster vector machines (SVMs), and various leveled cluster examination (HCA). Nectar characterization was accomplished by fractional slightest squares modeling (PLS). All PLS models created were able to precisely gauge the proper values of the parameters analyzed utilizing as input the voltammetric exploratory information (i.e., $r > 0.9$). This affirms the potential capacity of the VE-tongue for performing a quick characterization of honeys through PLS in which an uncomplicated, cost-effective test arrangement handle that does not require the utilize of extra chemicals is executed. Therefore, PLS and SVM are the perfect way">the most perfect way and offer assistance to preserve the quality as well electronic tongue as a progressed and the novel rising innovation has been effectively utilized for the fast recognizable proof of cocoa beans concurring with their topographical areas. Seven categories of cocoa beans from Ghana were utilized in this try. Electronic tongue framework was utilized for information securing, whereas three design acknowledgment strategies were connected comparatively to construct the segregation show. The exhibitions of the models were cross-validated to guarantee its steadiness. Exploratory comes about uncovering that Fisher's discriminant investigation (FDA) is superior than foremost component examination (PCA) for visualizing the cluster patterns. K-nearest neighbor (KNN) demonstrated was marginally superior than FDA showed at an ideal execution of 100 % within the preparing set and 98.8 % in the forecast set [38]. By and large, support vector machine (SVM) was prevalent in both FDA and KNN with 100 % segregation rate in both the preparing and expectation set at five PCs. The finding results in that electronic technology coupled with support vector machine technique can instantly, accurately, and reliable discriminate quality management assurance.

In [38], authors analyzed the impact of agriculture land conversion on regional food security using Normalized Difference Vegetation Index (NDVI) on Landsat Imagery using ML. Authors used satellite images and applied supervised learning, mainly Classification and Regression tree algorithm (CART) for classifying the areas into different regions according to the food availability in that region. The above mentioned is used for predicting the paddy field output considering the required factors like agriculture, land area, conversion rate, etc., the predicted values are compared with the growing demand of paddy field being calculated while assuming there will be no import or export of paddy, constant agriculture land conversion, and constant growing population. It was concluded that out of the four districts, Samigaluh district was labeled as food insecure, Districta , Girimulyo was labeled as food insecure,

Kalibawang was classified as food secure, and the last District Kokap was labeled as food inseure. In the year 2073, it is predicted that here will be food self-sufficiency in the region.

The authors compared six different ML algorithms named logistic regression, naive Bayes, SVM, neural network, k-nearest neighbor, and random forest to examine the food security status in the area of Punjab (Pakistan). To collect appropriate data in different first stage different zones based on the crop people growing there were selected, in the second stage homogeneity was considered for different crops mainly from 5 crops like sugarcane, maize, rice, wheat and cotton, in the later stage r 12 villages were selected and out of which 756 households were selected. Then an application was built and used, 756 households filled the data related to food intake of the previous seven days. When the data was collected, it was classified as training data and testing data. Later, the performance of each above-mentioned algorithm was calculated, out of all Random Forests stands out to be the best algorithm, and SVM provides the second best results for the data [39].

In [40], the authors used ML (machine learning) for famine prediction. Authors believed that early famine prediction can benefit in a large way, it can save many lives, it can give us time to get prepared for the upcoming famine. In this research, four machine learning techniques named K-nearest neighbor, SVM, Decision tree, Naïve bayes were used and the results were compared for four regions of Uganda called Northern, Southern, Eastern, Western. Out of these four region dataset from northern and easter region was the best for famine prediction. In the early stages twenty-four features like crop name, crop number, etc were used, but in the final stage nine out of these factors were not included as they were either transferred or reconstructed into other variables. Due to its high generalization performance, SVM along with KNN performs better in comparison to other techniques. Naïve Bayes was not able to perform very well due to the assumption of conditional independence among some variables. Decision tree gave the worst performance due to over fitting in the result. Authors also concluded that socio-economic data for specific areas and periods along with the satellite images needed to be used in combination to show the relationship in food security prediction.

In [41], the authors identified and classified the contamination present in the plant due to solutes and water which goes through roots of the plant. Utilizing cut-off, different parts of plants were examined for the fate of contaminants, later these cutoffs were used in algorithms for prediction of contamination. RCF (Root concentration factor) and FCF (fruit concentration factor) are measures of contaminants aggregated in edible tissues and roots of plants. To achieve their goal, the authors used ML algorithms like NN, k-means, and fuzzy logic. Different types of NN like MLP (Multi-Layer Perceptron), RCF (Radical Basis Function) were used for prediction of RCF (Root Contamination Factor) instead of limitation like in some cases instead of stopping at global maxima it may stop at local maxima. Out of these, MLP performed well and was selected for further use. Fuzzy logic was used to study the partitioning threshold of chemical compounds using physicochemical cutoffs and the simultaneous effects of compound properties. When using hierarchical clustering and partitional clustering methods to explore the datasets described

here, the k-means algorithm produced the most compact, isolated, and meaningful groups.

In the Internet of Things age, the meteorological department uses various types of sensors to determine humidity and temperature, among other things. MapReduce technology is used to efficiently analyze weather data using distributed algorithms. The advantage of using Map Reduce on Hadoop is that it can speed up data collection in an environment where the amount of data is growing by day. [43]

In the paper [42], authors evaluated the effectiveness of traditional handcrafted features and proposed using deep convolutional NN (ConvNets) for classifying of Fine-grained leaves. To explore the robustness of these features, the authors introduce a variety of conditions, such as translation, scaling, rotation, shading, and occlusion. In it, five types of algorithms were compared, named ConvNet (Deep Learning Neural Network), HCF (Hand Crafted Features), HoCS (Histogram of Curvature over Scale), and Combined (ConvNet features with HCFs) and the accuracy of all these were compared. The Combined approach achieves the highest level of accuracy with an average of 97.3% ±0.6% in comparison with traditional features, obtaining an average of 91.2% ± 1.6% compared to a typical mode. Here, Table 4 represents ML models in some recent research on demand prediction.

3. Proposed Methodology

Authors propose a new ecosystem for Agriculture Supply Chain (ASC) which will use ML for suggesting, predicting, managing, and fulfilling the demand for agriculture products. It will also decrease the financial load of the industries which will increase the income of farmers and it will provide knowledge of modern methods, technologies, rules, policies, and schemes made by the government for the betterment of farmers to the farmers.

Table 4. ML models for food security

ML Algorithm	Reference
[37] SVM, HCA	Emerging approach for analytical characterization and geographical classification of Moroccan and French honeys by means of a voltammetric electronic tongue
[38] KNN, SVM	Discrimination of cocoa beans according to geographical origin by electronic tongue and multivariate algorithms
[4] NDVI, CART	Identification of agricultural land use change based on machine learning for regional food security analysis in the mountainous region of Kulon Progo Regency.
[39] naive Bayes, k-nearest neighbor, logistic regression, SVM, neural network, RF	An Automatic Determining Food Security Status: Machine Learning based Analysis of Household Survey Data
[40] K-nearest neighbor, SVM, Decision tree, Naïve bayes	Machine learning classification technique for famine prediction
[41] NN, fuzzy logic, clustering algorithm	Examining plant uptake and translocation of emerging contaminants using machine learning: implications to food security
[42] ConvNet (Deep Learning Neural Network), HCF (Hand Crafted Features), HoCS	Evaluation of features for leaf classification in challenging conditions.

3.1 Present supply chain of products

In the present system, industry buys the whole crop at once and then processes it and later they supply it to the next level, it will take some and large capital which will lead industry people to borrow loan. The current system is given in Figure 1. Due to this loan, the price of the product increases and the same process is repeated on each level.

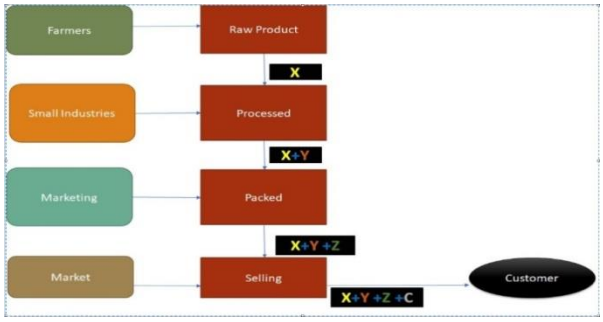

Figure 1. Current system

3.2 Proposed ecosystem for a supply chain of products

To make it more farmer, industry, and customer friendly and the authors propose a model where this whole system is controlled by centralised units and it will provide a token at each level of the system in exchange of service or raw products. These tokens will be converted to money once the product is sold, it will prevent the extra loan money, save time, and increase the efficiency. The proposed model is discussed in Figure 2.

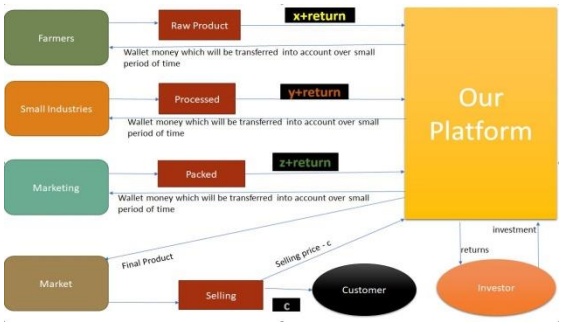

Figure 2. Proposed system

3.3 Use of Machine Learning Algorithms

We have divided the use of machine learning in 4 different clusters of crop production and for this project we are selecting only the most important ML process from each phase.

1. Data pre-processing- to use any type of ML algorithm, we feed it with a large amount of data and this step can deal with the processing of data and make it useable for training the data.

2. Model selection- in this step for each process of every cluster we need to select the best model

Cluster 1(Crop yield)- for crop yield there are many techniques like ANN, Bayesian network, Deep Learning, Decision tree, Ensemble Learning, Instance-Based Learning, SVM, Regression. The technique which will give the best result according to the conditions will be selected.

Cluster 2 (weather prediction)- for weather prediction there are many techniques like ANN, Deep Learning, Decision tree, Ensemble Learning, Instance-Based Learning, Regression. The technique which will give the best result according to the conditions will be selected.

Cluster 3(Demand Prediction, distribution management)- for demand prediction and distribution management, there are techniques like ANN, Genetic Algo, Clustering, Regression. The technique which will give the best result according to the conditions will be selected.

Cluster 4(Consumer Analysis)- for consumer analysis, there are many techniques like ANN Deep learning. The

technique which will give the best result according to the conditions will be selected.

3. Technique Usage- once the best technique is selected, it will be used in our model and once the user will give the input, the model will give an answer accordingly. For this problem, the following dataset (https://data.world/thatzprem/agriculture-india) is taken for result analysis.

The feature consists of the following: -

- State
- Year
- Season
- Crop
- Area
- Production
- Rainfall

Production being our dependent variable is our target variable. First of all, the authors have loaded the dataset using the pandas' package of Python and did some EDA (Exploratory Data Analysis), like what kind of data the authors have, and how the distribution of the data. Are there any null values?

Now, the authors had to convert the categorical data into numeric data because Machine Learning algorithms can only work on numerical data, therefore the authors have used Ordinal Encoding for the features such as StateSeasonCrop. Post that, authors divided the dataset into the train and test split in the 80:20 ratio. Since it is a regression problem the authors tried several ML algorithms like Linear Regression, Random Forest Regressor, Gradient Boosting Regressor. After experimenting with several parameters, the authors observed that the Gradient Boosting Algorithm was outperforming the other two in terms of the results, therefore we went ahead on fine tuning the model using hyperparameter tuning using various techniques like GridSearchCV and RandomizedSearchCV. Gradient boosting is a type of machine learning boosting. It relies on the intuition that the best possible next model, when combined with previous models, minimizes the overall prediction error. The key idea is to set the target outcomes for this next model to minimize the error. Since the authors used the tree-based algorithm, although due to differences in the scaling of the features, the authors did not do any standardization as well as normalization, as tree-based algorithms do not require scaling. As a part of the performance metrics, the authors have used Adjusted R-Squared values, which can be defined as a modified version of R-squared that has been adjusted for the number of predictors in the model.

4. Crop Recommendation

The dataset the authors have used can be found at

(https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset)

The feature consists of the following: -

- N - ratio of nitrogen content in soil
- P - ratio of phosphorus content in soil
- K - ratio of potassium content in soil
- temperature - temperature in degrees Celsius
- humidity - relative humidity in %
- ph - ph value of the soil
- rainfall - rainfall in mm

● Label - the target feature

Label being our dependent variable is our target variable which is denoting various kinds of crops. First of all, the authors have loaded the dataset using the panda's package of Python and did some EDA (Exploratory Data Analysis), like what kind of data the authors have, and how the distribution of the data. Are there any null values?

Now, the authors had to convert the categorical data into numeric data because Machine Learning algorithms can only work on numerical data, therefore the authors encoded the target variable label Post that authors divided the dataset into the train and test split in the 80:20 ratio. Since it is a classification problem we tried several ML algorithms like Logistic Regression, RandomForest Classifier, Support Vector Machine. Authors created a pipeline where we tested all these algorithms on our dataset and observed the results illustrated. Therefore, the authors observed that the RandomForest Classifier outperformed the other two by a significant margin. Random Forest is an ensemble algorithm which is also known as the bootstrap aggregation technique consisting of several trees which at the end give their individual score and the class which is in majority becomes the final predicted value. Since the authors used the tree-based algorithm, although due to differences in the scaling of the features, authors were not supposed to do any standardization as well as normalization, as tree-based algorithms do not require scaling, but due to the pipeline that the authors have built it consists of SVM as well as Logistic Regression model also which might get impacted due to scaling, therefore authors have done standardization on this dataset. As a part of performance metrics, the authors have used Confusion Metric which gives us several performance measures by looking at which can evaluate in this model.

5. Conclusion

The scope of ML (Machine Learning) in food demand and agriculture is very high. To meet the increasing demand of the population and ensuring the food security, the implementation of machine learning in agriculture and food chain management is a must. ML can be implemented in many stages of agriculture and security like Pre-Production phase (crop yield, soil properties), Production phase (weather prediction, Livestock Management), Processing Phase (Demand Management, Quality Management), Distribution (Transportation, Consumer analysis). Out of all applications of ML, image processing and data prediction are widely used. During the research, authors determine that out of all NN (neural networks), SVM are the most popular algorithms to be used. The food security can be ensured by taking care of crop during the production phase, using the images of the crop for preventing it from diseases, and ensuring the produced food reaches to the people who actually need it. The future scope of ML is very high, the present algorithms have many limitations which can be overcome by introducing more powerful algorithms.

References

[1]. Sharma, R., Kamble, S. S., Gunasekaran, A., Kumar, V., Kumar, A. (2020). A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. Computers & Operations Research, 119, 104926.

[2]. Khan, R., Dhingra, N., Bhati, N. (2022). Role of Artificial Intelligence in Agriculture: A Comparative Study. In Transforming Management with AI,

Big-Data, and IoT (pp. 73-83). Springer, Cham.

[3]. Van Klompenburg, T., Kassahun, A., Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709.

[4]. Sholikah, L. N., Nisa, Z. K., Pratama, B. F., Pradipta, A. G., Susanto, S., Prihanantya, A. S., Arif, S. S. (2021, November). Identification of agricultural land use change based on machine learning for regional food security analysis in the mountainous region of Kulon Progo regency. In IOP Conference Series: Earth and Environmental Science (Vol. 922, No. 1, p. 012060). IOP Publishing.

[5]. Jones, P., Salter, A. (2013). Modelling the economics of farm-based anaerobic digestion in a UK whole-farm context. Energy Policy, 62, 215-225.

[6]. Balezentiene, L., Streimikiene, D., Balezentis, T. (2013). Fuzzy decision support methodology for sustainable energy crop selection. Renewable and Sustainable Energy Reviews, 17, 83-93.

[7]. Passuello, A., Cadiach, O., Perez, Y., & Schuhmacher, M. (2012). A spatial multicriteria decision making tool to define the best agricultural areas for sewage sludge amendment. Environment international, 38(1), 1-9.

[8]. Patil, A., Kokate, S., Patil, P., Panpatil, V., & Sapkal, R. (2020). Crop Prediction using Machine Learning Algorithms. Kapila Journal of Research, 1(1), 1-8.

[9]. Bhende, M. J., and K. P. Kalirajan. "Technical efficiency of major food and cash crops in Karnataka (India)." Indian Journal of Agricultural Economics 62.902-2016-67386 (2007).

[10]. Agarwal, S., & Tarar, S. (2021). A hybrid approach for crop yield prediction using machine learning and deep learning algorithms. In Journal of Physics: Conference Series (Vol. 1714, No. 1, p. 012012). IOP Publishing.

[11]. Pudumalar, S., Ramanujam, E., Rajashree, R. H., Kavya, C., Kiruthika, T., & Nisha, J. (2017, January). Crop recommendation system for precision agriculture. In 2016 Eighth International Conference on Advanced Computing (ICoAC) (pp. 32-36). IEEE.

[12]. Nishant, P. S., Venkat, P. S., Avinash, B. L., & Jabber, B. (2020, June). Crop yield prediction based on Indian agriculture using machine learning. In 2020 International Conference for Emerging Technology (INCET) (pp. 1-4). IEEE.

[13]. Gao, H. (2021, January). Agricultural Soil Data Analysis Using Spatial Clustering Data Mining Techniques. In 2021 IEEE 13th International Conference on Computer Research and Development (ICCRD) (pp. 83-90). IEEE.

[14]. Agarkhed, J., Dattatraya, P. Y., & Patil, S. (2021). Precision agriculture with cluster-based optimal routing in wireless sensor network. International Journal of Communication Systems, 34(10), e4800.

[15]. Rasheed, N., Khan, S. A., Hassan, A., & Safdar, S. (2021). A Decision Support Framework for National Crop Production Planning. IEEE Access, 9, 133402-133415.

[16]. Ramírez-García, J., Carrillo, J. M., Ruiz, M., Alonso-Ayuso, M., & Quemada, M. (2015). Multicriteria decision analysis applied to cover crop species and cultivar selection. Field Crops Research, 175, 106-115.

[17]. Kumar, Y. J. N., Spandana, V., Vaishnavi, V. S., Neha, K., & Devi, V. G. R. R. (2020, June). Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 736-741). IEEE.

[18]. Alagurajan, M., & Vijayakumaran, C. (2020). ML Methods for Crop Yield Prediction and Estimation: An Exploration. International Journal of Engineering and Advanced Technology, 9(3).

[19]. Dahikar, S. S., & Rode, S. V. (2014). Agricultural crop yield prediction using artificial neural network approach. International journal of innovative research in electrical, electronics, instrumentation and control engineering, 2(1), 683-686.

[20]. Ramesh, D., & Vardhan, B. V. (2015). Analysis of crop yield prediction using data mining techniques. International Journal of research in engineering and technology, 4(1), 47-473.

[21]. Bendre, M. R., Thool, R. C., & Thool, V. R. (2016). Big data in precision agriculture through ICT: Rainfall prediction using neural network approach. In Proceedings of the international congress on information and communication technology (pp. 165-175). Springer, Singapore.

[22]. Abdel-Kader, H., Abd-El Salam, M., & Mohamed, M. (2021). Hybrid machine learning model for rainfall forecasting. Journal of Intelligent Systems and Internet of Things, 1(1), 5-12.

[23]. Jain, S., & Ramesh, D. (2020, February). Machine Learning convergence for weather based crop selection. In 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) (pp. 1-6). IEEE.

[24]. Zhang, L., Huang, Z., Liu, W., Guo, Z., & Zhang, Z. (2021). Weather radar echo prediction method based on convolution neural network and Long Short-Term memory networks for sustainable e-agriculture. Journal of Cleaner Production, 298, 126776.

[25]. Kumar, S., Chaddha, H., Kumar, H., & Singh, R. K. (2019). Price forecasting system for crops at the time of sowing. Int Res J Eng Technol (IRJET), 6(4).

[26]. Paterakis, N. G., Mocanu, E., Gibescu, M., Stappers, B., & van Alst, W. (2017, September). Deep learning versus traditional machine learning methods for aggregated energy demand prediction. In 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe) (pp. 1-6). IEEE.

[27]. Samuel, P., Sahithi, B., Saheli, T., Ramanika, D., & Kumar, N. A. (2020). Crop price prediction system using machine learning algorithms. Quest Journals Journal of Software Engineering and Simulation.

[28]. Nguyen, H. V., Naeem, M. A., Wichitaksorn, N., & Pears, R. (2019). A smart system for short-term price prediction using time series models. Computers & Electrical Engineering, 76, 339-352.

[29]. Khan, R., Tyagi, N., & Chauhan, N. (2021). Safety of food and food warehouse using VIBHISHAN. Journal of Food Quality, 2021.

[30]. Cooray, S. E., Tassé, M. J., Barnhill, J., & Bhaumik, S. (2022). Anxiety and Stress-Related Disorders in People with Intellectual Disability/Disorders of Intellectual Development. In Textbook of Psychiatry for Intellectual Disability and Autism Spectrum Disorder (pp. 583-608). Springer, Cham.

[31]. Georgeou, N., Hawksley, C., Wali, N., Lountain, S., Rowe, E., West, C., & Barratt, L. (2022). Food security and small holder farming in Pacific Island countries and territories: A scoping review. PLOS Sustainability and Transformation, 1(4), e0000009.

[32]. Chen, Y., & Li, Y. (2019). Intelligent autonomous pollination for future farming-a micro air vehicle conceptual framework with artificial intelligence and human-in-the-loop. IEEE Access, 7, 119706-119717.

[33]. De Croon, G. C. H. E., De Clercq, K. M. E., Ruijsink, R., Remes, B., & De Wagter, C. (2009). Design, aerodynamics, and vision-based control of the DelFly. International Journal of Micro Air Vehicles, 1(2), 71-97.

[34]. Ma, K. Y., Chirarattananon, P., Fuller, S. B., & Wood, R. J. (2013). Controlled flight of a biologically inspired, insect-scale robot. Science, 340(6132), 603-607.

[35]. Y. Chen and Y. Li, Computational Intelligence Assisted Design (In Industrial Revolution 4.0). Boca Raton, FL, USA: CRC Press, 2018. [Online]. Available: https://www.taylorfrancis.com/books/9 7814987 60676

[36]. Chen, Y., Song, Z., Zhang, G., Majeed, M. T., & Li, Y. (2018). Spatio-temporal evolutionary analysis of the township enterprises of Beijing suburbs using computational intelligence assisted design framework. Palgrave Communications, 4(1), 1-14.

[37]. El Hassani, N. E. A., Tahri, K., Llobet, E., Bouchikhi, B., Errachid, A., Zine, N., & El Bari, N. (2018). Emerging approach for analytical characterization and geographical classification of Moroccan and French honeys by means of a voltammetric electronic tongue. Food chemistry, 243, 36-42.

[38]. Teye, E., Huang, X., Han, F., & Botchway, F. (2014). Discrimination of cocoa beans according to geographical origin by electronic tongue and multivariate algorithms. Food Analytical Methods, 7(2), 360-365.

[39]. Razzaq, A., Ahmed, U. I., Hashim, S., Hussain, A., Qadri, S., Ullah, S., ... & Asghar, A. (2021). An Automatic Determining Food Security Status: Machine Learning based Analysis of Household Survey Data. International Journal of Food Properties, 24(1), 726-736.

[40]. Okori, W., & Obua, J. (2011, July). Machine learning classification technique for famine prediction. In Proceedings of the world congress on engineering (Vol. 2, No. 1, pp. 4-9).

[41]. Bagheri, M., Al-Jabery, K., Wunsch, D., & Burken, J. G. (2020). Examining plant uptake and translocation of emerging contaminants using machine learning: Implications to food security. Science of the total environment, 698, 133999.

[42]. Hall, D., McCool, C., Dayoub, F., Sunderhauf, N., & Upcroft, B. (2015, January). Evaluation of features for leaf classification in challenging conditions. In 2015 IEEE Winter Conference on Applications of Computer Vision (pp. 797-804). IEEE.

[43]. Phasinam, Khongdet, Thanwamas Kassanuk, and Mohammad Shabaz. "Applicability of internet of things in smart farming." Journal of Food Quality 2022 (2022).