

# Crop Yield And Its Nutrition Prediction Using Machine Learning A Regression Algorithm

<sup>1</sup>**Akundi Sai Hanuman**, Professor of CSE , Gokaraju Lailavathi womens engineering, Hyderabad [admnglwec@gmail.com](mailto:admnglwec@gmail.com)

<sup>2</sup>**R. Sowjanya**, Student, Dept of CSE, Gokaraju Lailavathi womens engineering, Hyderabad

**ABSTRACT:** Agriculture is backbone about India's economy because it supports over half about country's population. continued success about agriculture is now seriously threatened through changes in weather, climate, & other environmental factors. choice help device for Crop Yield and nutrition Prediction (CYP), which remembers helping choices for which yields towards plant & what towards do during development time about yields, is where AI (ML) assumes an imperative part. CYP traits are primary focus about current study, which is organized around a systematic review. Likewise, various procedures have been made towards break down agrarian creation forecast using computer based intelligence strategies. main disadvantages about neural network are a lower accuracy in crop yield prediction & a lower relative error. Like this, managed learning calculations neglected towards perceive nonlinear connection among info & result factors, which introduced a test during choice, reviewing, or arranging about natural products. Numerous studies were encouraged towards establish an accurate & efficient model for crop classification, including crop yield estimation based on soil, weather, & crop disease, crop classification based on growing phase, & so forth. accuracy about various machine learning (ML) methods used towards estimate crop yields is thoroughly examined in aforementioned study.

**Keywords** – Machine learning, crop yield prediction.

## 1. INTRODUCTION

Agriculture has always been most important activity in every society & civilization certain has ever existed. Not only does it play a significant role in expanding economy, but it is also essential towards our survival. Additionally, it is a crucial sector for future about humanity & Indian economy. Additionally, it accounts for a significant portion about employment. demand for production has significantly increased over time. In order towards produce in large quantities, people use technology in completely inappropriate ways. Consistently, new crossover sorts are made. These sorts, in any case, don't offer similar basic parts as a yield developed normally. soil is degraded through these synthetic methods. Environmental devastation grows as a result about everything. majority about these unconventional strategies are utilized towards avoid losses. However, when agricultural growers have access towards accurate crop production data, loss decreases. AI is a quickly growing procedure certain upholds decision-production across all ventures towards give most helpful about its applications. Models can be examined prior towards deployment, which is beneficial for majority about modern devices. basic idea is towards use models from machine learning towards increase agricultural sector

throughput. How much data gave all through preparation period, as there were more boundaries in general, is another variable certain impacts forecast. Accuracy farming, which focuses on quality over negative natural factors, would be fundamental concentration. Several classifiers for machine learning, such as Logistic Regression, Naive Bayes, Random Forest, & others, are used towards point towards a pattern so certain accurate predictions can be made & erratic patterns about temperature & rainfall can be taken into account. Our examination about previously mentioned AI classifiers drove us towards end certain Arbitrary Woods technique offers most elevated level about exactness. system uses historical data towards make crop predictions. Historical temperature, weather, & other variables are used towards provide information. Our application runs calculation & presentations a rundown about harvests certain match inputted information & their expected yield esteem.

Agriculture has been practiced in India since Indus Valley Civilization. In aforementioned industry, India is positioned second. agricultural & related industries account for 15.4% about GDP (gross domestic product), & they employ approximately 31% about all workers. Net cropped area is dominated through India, followed through United States & China. biggest financial area as far as populace variety is horticulture, which means a lot towards India's whole financial construction. Agriculture's economic contribution towards India's GDP is decreasing in tandem with country's overall economic expansion as a result about industrialization revolution. use about machine learning techniques towards predict crop

productivity is at heart about problem statement. project's goal is towards help users choose best crop for production in order towards maximize output & profits. proposed framework makes forecasts through breaking down information while endeavoring towards address weaknesses about current techniques. We propose expanding variety about crops certain can be grown throughout season & creating a system certain takes into account factors certain have greatest impact on how well a crop grows. aforementioned would make it simpler for farmers towards select crops certain will yield a high yield and, as a result, maximize income, lowering suicide rate. framework incorporates a module for yield expectation. aforementioned module provides user with two choices: it is possible certain they can pick a particular harvest & view its yield, or they might see best 5 harvests generally speaking regarding yield.

## 2. LITERATURE REVIEW

The writers about article Anticipating Yield about harvest yield utilizing AI calculations are P. Priya, U. Muthaiah, & M. Balamurugan[1]. country's economy is towards a great extent driven through rural area. Environmental change & other natural changes are turning into a serious risk towards farming. A significant technique for finding practical & productive answers for aforementioned issue is AI (ML). Crop yield forecast is method involved with assessing harvest's creation in view about authentic information, like climate, soil, & past harvest yields. aforementioned study utilizes Arbitrary Woodland calculation towards estimate farming yield in light about accessible information. models were built

utilizing genuine information from Tamilnadu, & they were tried utilizing tests. Prior towards developing in a rural field, rancher could utilize forecast towards assess harvest's creation. Irregular Woodland, one about most impressive & popular directed AI calculations, is utilized towards foresee agrarian efficiency in future appropriately.

The journalists about title Uses about AI methods in agrarian yield creation are Mishra. S, Mishra. D, & Santra. G. H [2]. aforementioned article was written with an end goal towards reconsider research papers on relevance about AI techniques towards creation about agrarian yields. Measurable Investigation/Strategies: aforementioned system is fresh out about box new for overseeing farming yield creation. For directorate about financial matters & insights towards give huge approach decisions like import-send out, cost showcasing dispersion, & so on, precise & opportune harvest creation projections are required. In any case, one should take note about certain these prior projections are not objective projections since they require broad spellbinding assessment in light about a large number about subjective components. In aforementioned way, it is important towards build objective harvest creation expectations certain are measurably solid. Those headways in figuring & data stockpiling have made a lot about information accessible. Discoveries: Because about intricacy about information got from aforementioned crude information, new methodologies & strategies, such AI, have been created towards incorporate comprehension about information with crop yield assessment. aforementioned study expects towards evaluate these

state about art techniques so huge connections between various factors in information base can be found.

Crop yield expectation has been carried out through Balamurugan[3] utilizing just arbitrary timberland classifier. towards expect agrarian result, different elements like precipitation, temperature, & season were thought of. On datasets, no further AI techniques were utilized. Since elective calculations were inadequate with regards to, correlation & measurement wasn't possible, making it difficult towards give best calculation.

AI calculations for crop yield expectation & nitrogen status evaluation in accuracy horticulture: A survey, Anna Chlingaryana, Salah Sukkarieha, Brett Whelan, et al., [4], PCs & Gadgets in Farming, vol. essential supporter about public economy is horticultural result. Notwithstanding, factors like precipitation, temperature, moistness, & other natural variables altogether affect farming efficiency. Early harvest yield anticipating can assist ranchers with dealing with their harvests all more actually. Choice Tree calculation, an AI strategy, is utilized in aforementioned paper towards impersonate various informational collections. informational collection is exposed towards calculation, & results have been thought about.

A.T.M Shakil Ahamed, Kallal Das, Mohammad Tanzir Kabir, Nazmul Hossain, Navid Tanzeem Mahmood, & Faridur Rahmanetal.,[5] Applying Information Mining Strategies towards Anticipate Yearly Yield about Significant Harvests & Recommend Establishing Various Harvests in

Various Regions in Bangladesh. Horticultural yield is impacted through various elements, including science, environment, economy, & topography. A few factors affect farming in different ways, & these impacts can be estimated with right factual methods. It is feasible towards acquire data or information certain can assist ranchers & government associations with pursuing better choices & spread out methodologies certain outcome in more noteworthy creation through utilizing such strategies & procedures towards verifiable yield about harvests. Our essential concern is utilization about information mining procedures towards agrarian information towards separate data & gauge crop creation for significant oat crops in critical regions about Bangladesh.

### 3. METHODOLOGY

Neural networks were utilized in majority about successful models; KNN regression methods for CYP; furthermore, for best expectation, various ML strategies were likewise applied. problems certain crop yield prediction research using machine learning currently faces are listed below:

1. development, maintenance, & repair about ML algorithms required significant expenditures due towards their complexity.
2. For crop yield prediction (wheat, mustard), ML method combined input & output data, but no statistically improved results were obtained.
3. Due towards nature about linear link in parameters, regression model was unable towards accurately

forecast complex situations with extreme value data & nonlinear data.

4. presentation about K-NN models certain were at certain point being used for characterization & yield expectation was diminished in view about nonlinear & profoundly versatile KNN challenges. They were worked in a territory model certain expanded info vector's dimensionality & created characterization turmoil.

5. A suitable classification decision was not made because there were insufficient data towards estimate crop production.

#### Disadvantages:

Nonlinear & highly adaptive KNN difficulties decreased classification & yield prediction performance about existing K-NN models.

#### PROPOSED SYSTEM

The viable use about AI procedures & its evaluation are primary subjects about aforementioned undertaking. work presented here also takes into account erratic data from temperature & rainfall records in order towards obtain a consistent trend. Crop yield predictions are made through taking into account all about factors, as opposed towards usual practice about only considering one aspect at a time.

#### Advantages:

- models presented in aforementioned study take into account factors like temperature & weather conditions towards provide accurate crop yield

forecasts.

- calculations used Choice tree & Arbitrary woodland are straightforward & fruitful in forecast.
- In end, experimental study demonstrated how ML can be applied towards agricultural domain, such as soil characteristics, in addition towards current research, towards improve crop prediction.

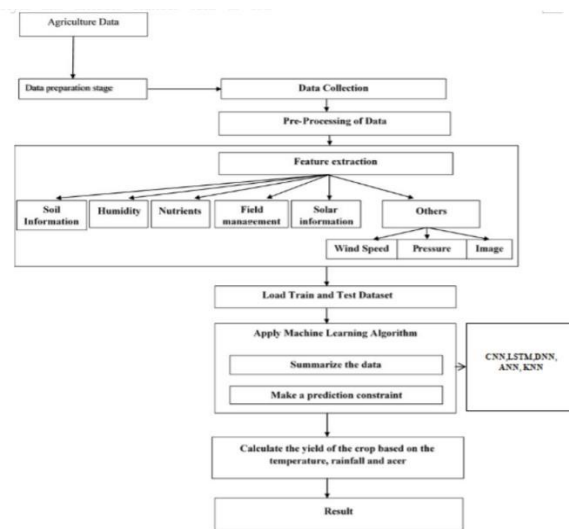


Fig.2: System architecture

### MODULES:

Following modules have been created for aforementioned project.

Data cleansing & validation (Module1)

Data exploration, visualisation analysis, & dummying (Module 2)

Training a model using Random Forest technique & provided attributes (Module 3).

Random Forest, KNN, & Decision Tree Performance Metrics (Module-4)

Measures about linear regression's performance (Module-05)

## 4. IMPLEMENTATION

Random forests, also known as random decision forests, are a group learning method for arrangement, relapse, & different undertakings. At hour about preparing, countless choice trees are developed, & class certain addresses method about classes (relapse) or mean forecast (grouping) about each tree is yield. Unpredictable decision woods ideal for decision trees' affinity for over fitting towards their readiness set. Unpredictable forest is a sort about directed artificial intelligence estimation considering gathering learning. Gathering learning is a kind about learning wherein different occurrences about similar calculation or different calculations are consolidated towards make a more intense forecast model. expression "irregular backwoods" comes from way certain arbitrary woodland calculation makes a timberland about trees through consolidating different calculations about a similar sort, for example, numerous choice trees. Relapse & arrangement assignments can both be achieved with assistance about irregular woods calculation. Coming up next are key advances drew in with playing out erratic forest estimation pick N unpredictable records from dataset. In view about these N records, build a choice tree. Rehash stages 1 & 2 for quantity about trees you need in your calculation. For each new record in a relapse issue, each tree in woodland predicts an incentive for Y(output). last worth not set

in stone through taking normal about each about qualities anticipated through every one about trees in woodland. Then again, in case about a portrayal issue, each tree in forest area predicts class towards which new record has a spot. new record is then positioned in classification certain got greater part vote.

KNN: K-Nearest Neighbor is a controlled man-made intelligence estimation which stores all cases contrast with getting ready information about interest in n-layered space. For genuine esteemed information, it returns mean about k closest neighbors, while for obscure discrete information, it examines nearest k number about saved occurrences (closest neighbors) & returns most widely recognized class as expectation. Utilizing accompanying question, commitment about each about neighbors is weighted as per their distance somewhere far off weighted closest neighbor calculation, with nearest neighbors getting more noteworthy weight. Since it midpoints k-closest neighbors, KNN ordinarily functions admirably with uproarious information. managed k-closest neighbors calculation is an order calculation: it takes a ton about checked concentrations & uses them towards sort out some way towards name different core interests. It takes a gander at named focuses nearest towards new point — those are new point's neighbors — & requests certain those neighbors vote on which mark ought towards be applied towards new point (the "k" demonstrates quantity about neighbors it checks). utilizes whole preparation set towards make forecasts about approval set. KNN makes an assumption regarding another model through means about glancing through

entire set towards consider towards be k "closest" cases.

Decision Tree: It is perhaps about most great & popular estimation. decision tree calculation is an illustration about a managed learning calculation. It 24 is powerful for both clear cut & ceaseless result factors. Premises about Choice Tree: Around beginning, we consider whole arrangement set as root. Credits are believed towards be outright for information gain, attributes are believed towards be diligent. Records are conveyed recursively in view about property estimations. Ascribes are arranged through root or interior hub utilizing measurable strategies. construction about a choice tree is utilized towards make grouping or relapse models. A choice tree is at same time created gradually while an informational collection is separated into increasingly small subsets. A decision center has something like two branches & a leaf center point tends towards a portrayal or decision. top most decision center point in a tree which looks at towards best marker called root center. Choice trees can manage mathematical & straight out information. construction about a choice tree is utilized towards make grouping or relapse models. For grouping, it utilizes a thorough & fundamentally unrelated on off chance certain standard set. preparation information are utilized towards get familiar with standards each in turn in a successive way. tuples covered through guidelines are taken out each time a standard is learned. On preparation set, aforementioned strategy go on until an end condition is met. Partition & overcome starting from top is utilized towards assemble it. All about qualities should be straight out. If not, they

ought towards be discretized in advance. Utilizing data gain idea, credits at highest point about tree are recognized on grounds certain they greaterly affect order. A choice tree might reflect oddities because about commotion or exceptions since it is not difficult towards over-fit it, bringing about an excessive number about branches.

Linear Regression: For prescient examination, linear regression is a tranquil & clear factual relapse method certain shows connection between constant factors. expression "straight relapse" alludes towards strategy certain exhibits direct relationship certain exists between autonomous variable (the X-hub) & reliant variable (the Y-hub). If there is a lone data variable (x), such straight backslide is called fundamental direct backslide. Also, aforementioned sort about direct relapse is alluded towards as numerous straight relapse on off chance certain there are more than one information variable. connection between factors is portrayed through a slanted straight line in direct relapse model.

**5. EXPERIMENTAL RESULTS**

	Algorithms	Accuracy	Standard Deviation
0	Random Forest	96.67	0.691015
1	Decision-tree	92.43	2.520343
2	KNN Classifier	98.18	0.668450

Fig.2: Output

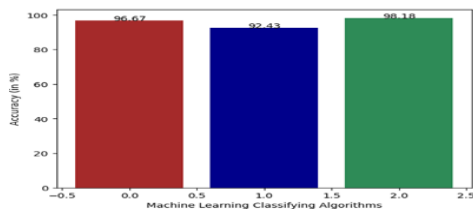


Fig.3: Output

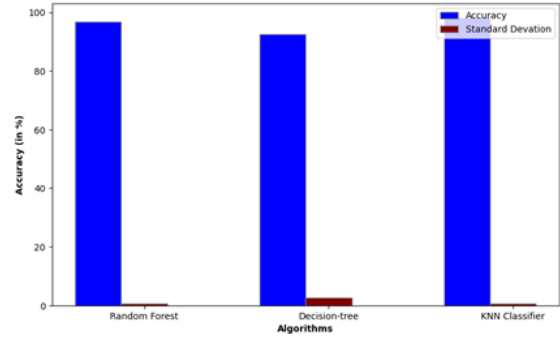


Fig.4: Output

**6. CONCLUSION**

The scope about elements shrouded in flow research exertion are for most part founded on accessibility about information, & each exploration will analyze CYP using ML strategies certain are unmistakable from highlights. selection about features was largely influenced through availability about data collection & was based on geological position, scale, & crop features. However, better outcomes were not always achieved through using more features. Consequently, testing was conducted towards identify few features with highest performance certain were also utilized in studies. In end, experimental investigation demonstrated how crop prediction can be improved through combining ML & agricultural domain. In terms about how temperature changes affect agriculture, however, there was still room for improvement. delay in bordering topographical areas, for instance, necessitated additional explicit treatment & should be primary focus about subsequent investigations. Finally, nonparametric portion about model is completed with features from deterministic crop models in order towards provide perfect

statistical CO<sub>2</sub> fertilization. Further exploration would further develop farming yield assessment through sticking towards previously mentioned objectives. We should develop & plan a model in view about DL for CYP relying upon consequences about examination. recommended framework depends on Irregular Backwoods calculation, which produces various harvest yields. Arbitrary woods utilizes recommended framework's precision towards deliver results certain are exact towards inside 96-97%. primary focus about aforementioned research is application about machine learning algorithms towards crop forecasting & yield estimation. computation about exactness utilizes an assortment about AI strategies. Random Forest classifier was used towards make crop predictions based on selected data. developed a method for anticipating crops using data from past. recommended strategy helps ranchers in picking which yield towards establish in field. purpose about aforementioned activity is towards learn more about crop, which can be used towards produce a harvest certain is both productive & efficient.

## 7. FUTURE SCOPE

In coming years, you might want towards use a data-independent system. Regardless about configuration, our framework should work with a similar precision. Incorporating soil data into framework is favorable since crop choice likewise thinks about soil information as an element. Crop development likewise requires compelling water system. Rainfall might indicate whether or not there is a need for more water. It is possible towards elevate aforementioned

study towards a higher level through making it accessible towards all about India.

## REFERENCES

- [1] Kailasam, S., Achanta, S.D.M., Rama Koteswara Rao, P., Vatambeti, R., Kayam, S. (2022). An IoT-based agriculture maintenance using pervasive computing with machine learning technique. *International Journal of Intelligent Computing and Cybernetics*, 15(2), pp. 184–197.
- [2] Ramaiah, V. S., Singh, B., Raju, A. R., Reddy, G. N., Saikumar, K., & Ratnayake, D. (2021, March). Teaching and Learning based 5G cognitive radio application for future application. In 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) (pp. 31-36). IEEE.
- [3] Mohammad, M. N., Kumari, C. U., Murthy, A. S. D., Jagan, B. O. L., & Saikumar, K. (2021). Implementation of online and offline product selection system using FCNN deep learning: Product analysis. *Materials Today: Proceedings*, 45, 2171-2178.
- [4] Padmini, G. R., Rajesh, O., Raghu, K., Sree, N. M., & Apurva, C. (2021, March). Design and Analysis of 8-bit ripple Carry Adder using nine Transistor Full Adder. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1982-1987). IEEE.
- [5] Dr. k. Raju, A. Sampath Dakshina Murthy, Dr. B. Chinna Rao, Sindhura Bhargavi, G. Jagga Rao, K. Madhu, K. Saikumar. A Robust And Accurate



- Video Watermarking System Based On SVD Hybridation For Performance Assessment International Journal of Engineering Trends and Technology, 68(7),19-24.
- [6] Saba, S. S., Sreelakshmi, D., Kumar, P. S., Kumar, K. S., & Saba, S. R. (2020). Logistic regression machine learning algorithm on MRI brain image for fast and accurate diagnosis. International Journal of Scientific and Technology Research, 9(3), 7076-7081.
- [7] Saikumar, K. (2020). RajeshV. Coronary blockage of artery for Heart diagnosis with DT Artificial Intelligence Algorithm. Int J Res Pharma Sci, 11(1), 471-479.
- [8] Saikumar, K., Rajesh, V. (2020). A novel implementation heart diagnosis system based on random forest machine learning technique International Journal of Pharmaceutical Research 12, pp. 3904-3916.
- [9] Raju K., Chinna Rao B., Saikumar K., Lakshman Pratap N. (2022) An Optimal Hybrid Solution to Local and Global Facial Recognition Through Machine Learning. In: Kumar P., Obaid A.J., Cengiz K., Khanna A., Balas V.E. (eds) A Fusion of Artificial Intelligence and Internet of Things for Emerging Cyber Systems. Intelligent Systems Reference Library, vol 210. Springer, Cham. [https://doi.org/10.1007/978-3-030-76653-5\\_11](https://doi.org/10.1007/978-3-030-76653-5_11)
- [10] Sankara Babu B., Nalajala S., Sarada K., Muniraju Naidu V., Yamsani N., Saikumar K. (2022) Machine Learning Based Online Handwritten Telugu Letters Recognition for Different Domains. In: Kumar P., Obaid A.J., Cengiz K., Khanna A., Balas V.E. (eds) A Fusion of Artificial Intelligence and Internet of Things for Emerging Cyber Systems. Intelligent Systems Reference Library, vol 210. Springer, Cham. [https://doi.org/10.1007/978-3-030-76653-5\\_12](https://doi.org/10.1007/978-3-030-76653-5_12)
- [11] Kiran Kumar M., Kranthi Kumar S., Kalpana E., Srikanth D., Saikumar K. (2022) A Novel Implementation of Linux Based Android Platform for Client and Server. In: Kumar P., Obaid A.J., Cengiz K., Khanna A., Balas V.E. (eds) A Fusion of Artificial Intelligence and Internet of Things for Emerging Cyber Systems. Intelligent Systems Reference Library, vol 210. Springer, Cham. [https://doi.org/10.1007/978-3-030-76653-5\\_8](https://doi.org/10.1007/978-3-030-76653-5_8)
- [12] Shravani, C., Krishna, G. R., Bollam, H. L., Vatambeti, R., & Saikumar, K. (2022, January). A Novel Approach for Implementing Conventional LBIST by High Execution Microprocessors. In 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 804-809). IEEE.
- [13] Nagendram, S., Nag, M. S. R. K., Ahammad, S. H., Satish, K., & Saikumar, K. (2022, January). Analysis For The System Recommended Books That Are Fetched From The Available Dataset. In 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 1801-1804). IEEE.
- [14] Jothisna, V., Patel, I., Raghu, K., Jahnavi, P., Reddy, K. N., & Saikumar, K. (2021, March). A Fuzzy Expert System for The Drowsiness Detection from Blink Characteristics. In 2021 7th International Conference on Advanced Computing and

- Communication Systems (ICACCS) (Vol. 1, pp. 1976-1981). IEEE.
- [15] Appalaraju, V., Rajesh, V., Saikumar, K., Sabitha, P., & Kiran, K. R. (2021, December). Design and Development of Intelligent Voice Personal Assistant using Python. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 1650-1654). IEEE.
- [16] Naidu, T. P., Gopal, K. A., Ahmed, S. R., Revathi, R., Ahammad, S. H., Rajesh, V., ... & Saikumar, K. (2021, December). A Hybridized Model for the Prediction of Heart Disease using ML Algorithms. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 256-261). IEEE.
- [17] Kailasam, S., Achanta, S. D. M., Rao, P. R. K., Vatambeti, R., & Kayam, S. (2021). An IoT-based agriculture maintenance using pervasive computing with machine learning technique. International Journal of Intelligent Computing and Cybernetics.
- [18] Koppula, N., Sarada, K., Patel, I., Aamani, R., & Saikumar, K. (2021). Identification and Recognition of Speaker Voice Using a Neural Network-Based Algorithm: Deep Learning. In Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies (pp. 278-289). IGI Global.
- [19] Rao, K. S., Reddy, B. V., Sarada, K., & Saikumar, K. (2021). A Sequential Data Mining Technique for Identification of Fault Zone Using FACTS-Based Transmission. In Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies (pp. 408-419). IGI Global.