International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 08 | August - 2022 Impact Factor: 7.185 ISSN: 2582-3930

CROP YIELD PREDICTION USING MACHINELEARNING MODELS

N.Kesava Rao ,D.Ramya Sree, Ch.Mohan Rupa, Ch.Pavani, Ch.Sowdamini
Narayana Engineering College, Gudur

Abstract – This task predicts the yield of practically a wide range of harvests that are planted in India. This content makes novel by the utilization of basic boundaries like State, locale, season, region and the client can anticipate the yield of the harvest where year the person needs to. The paper utilizes different Machine learning models procedures like Random backwoods, Decision tree, KNN, SVM, and Navie Bayes supporting calculations to anticipate the yield and uses the idea for improving the calculations to give a superior forecast.

Index Terms – crop yield prediction, machine learning, random forest.

I. INTRODUCTION

In our assessment, which we found in the past investigation papers is that everyone uses climatic components like precipitation, sunlight and plant factors like soil type, supplements moved by the soil (Nitrogen, Potassium, etc) but the issue is we truly need to collect the data and a while later an untouchable does this assumption and subsequently it is explained for the farmer and this requires a lot of effort for the farmer and he doesn't sort out the science behind these components. To work on it and which can be directly used by the farmer this paper uses fundamental components like which state and area is the farmer from, which crop and in what season (as in Kharif, Rabi, etc.). In India, there are more than 100 yields laid out around the whole country. These yields are requested for better appreciation and portrayal.

The data for this assessment has been gotten from the Indian Government Repository [1]. The data involves properties - State, District, Crop, Season, Year, Area and Production with around 2.5 Lakh insights.

II. LITERATURE SURVEY

A. An improved crop yield prediction model using bee hive clustering approach for agricultural data sets

Rural specialists over the world demand the requirement for an effective instrument to anticipate and further develop the harvest development. The requirement for an incorporated harvest development control with exact prescient yield the executives procedure is profoundly felt among

cultivating local area. The intricacy of anticipating the harvest yield is exceptionally due to complex variable measurements and inaccessibility of prescient displaying approach, which prompts misfortune in crop yield. This examination paper proposes a harvest yield forecast model (CRY) which deals with a versatile group approach over powerfully refreshed verifiable yield informational index to foresee the harvest yield and further develop the dynamic in accuracy horticulture. CRY utilizes apiary displaying way to deal with investigate and arrange the harvest in light of yield development design, yield. CRY characterized dataset had been tried utilizing Clementine over existing yield space information.

B. An intelligent system based on kernel methods for crop yield prediction

This paper presents work on fostering a product framework for foreseeing crop yield from environment and ranch information. At the center of this framework is a strategy for solo parceling of information for finding spatio-transient examples in environment information utilizing piece techniques which offer solidarity to manage complex information. For this reason, a powerful weighted part k-implies calculation integrating spatial limitations is introduced. The calculation can actually deal with commotion, anomalies and auto-relationship in the spatial information, for powerful and productive information examination, and in this manner can be utilized for anticipating oil-palm yield by dissecting different elements influencing the yield of crop.

C. A Fuzzy Logic based Crop Yield Prediction using Temperature and Rainfall Parameters Predicted through ARMA, SARIMA, and ARMAX models.

Farming assumes a critical part in the economy of India. This makes crop yield expectation a significant undertaking to assist with helping India's development. Crops are delicate to different climate peculiarities like temperature and precipitation. Thusly, it becomes pivotal to incorporate these elements while foreseeing the yield of a harvest. Weather conditions determining is a confounded cycle. In this work, three strategies are utilized to conjecture ARMA (Auto Regressive Moving Average), SARIMA (Seasonal Auto Regressive Integrated Moving Average) and ARMAX (ARMA with exogenous factors). The exhibition of the three is looked at and the best model is utilized to foresee precipitation and temperature which are thus used to anticipate the harvest yield in light of a fluffy rationale model.

D. Crop Yield Prediction Using Data Analytics and Hybrid Approach

Rural information is being created continually and enourmosly. Accordingly, farming information has come in the period of enormous information. Brilliant advances contribute in information assortment utilizing electronic gadgets. In our undertaking we will break down and mine this rural information to obtain helpful outcomes utilizing advances like information examination and AI and this outcome will be given to ranchers for better harvest yield concerning proficiency and efficiency.

III. PROPOSED WORK

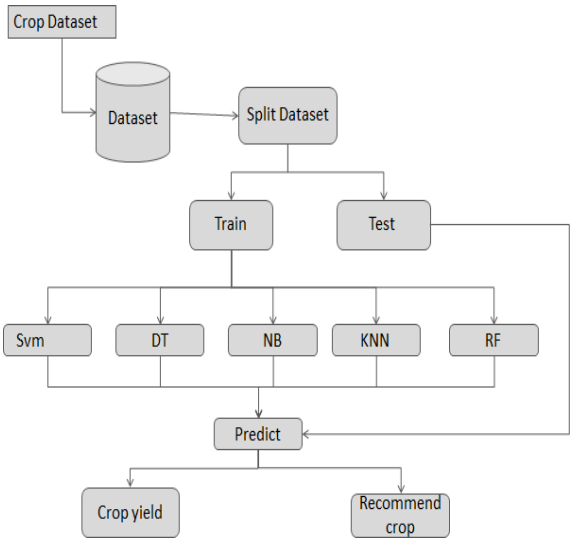

Fig. 1: System Workflow

The above diagram shows the crop yield prediction and analysis of the our proposed system using different machine learning models.

Implementation Modules

Preprocessing

For the given informational collection, there are many 'NA' values which are separated in python. Moreover, as the informational collection comprises of numeric information, we utilized hearty scaling, which is very like standardization, yet it rather utilizes the interquartile range though standardization is something which standardization recoils the information regarding zero to one.

Prediction

In this module, to foresee the harvest yield and creation in light of state name, crop name, no. of sections of land, and soil type utilizing different AI Classifier. And furthermore suggest the harvest in light of state and soil type.

Graphical Analysis

In this period of the Implementation client can get the reasonable picture examination of the reason for death investigation. Different variables think about for the diagram examination. In this stage plot the outlines like pie diagram, bar graph thus others.

Implementation Algorithm

This venture involved order method for expectation. Calculation has been chosen by assessing each regulated AI method. The objective of grouping is to anticipate future occasion by every classifier. In this work four classifiers are utilized in particular Random Forest, Decision Tree Regression Navie Bayes , KNN and Svm.The

forecast aftereffect of all classifiers are broke down and looked at.

Random Forest

- It produces multi choice trees from which every choice tree utilizes a piece of information test and predicts the outcome.
- Then, at that point, the outcome which was accomplished by greatest number of trees is considered as the last forecast.
- Irregular backwoods is a Supervised Learning calculation which utilizes outfit learning technique for grouping and relapse. Irregular woods is a packing procedure and the trees in irregular woodlands run in lined up with practically no communications.
- A Random Forest works by developing a few choice trees during preparing time and yielding the mean of the classes as the forecast of the relative multitude of trees.

Decision Tree

- Trees are developed through an algorithmic methodology that recognizes ways of dividing the informational index in view of various circumstances.
- It is one of the most generally involved useful strategies for directed learning.
- These are non-parametric technique utilized for both arrangement and relapse.

K Nearest Neighbour

- K-Nearest Neighbor is one of the easiest Machine Learning calculations in view of Supervised Learning strategy.
- K-NN calculation stores every one of the accessible information and characterizes another information point in light of the closeness. This implies when new information shows up then it tends to be effectively grouped into a well suite classification by utilizing K-NN calculation.
- K-NN calculation can be utilized for Regression as well with respect to Classification however for the most part it is utilized for the Classification problems.

Support Vector Machine

- Support Vector Machine or SVM is one of the most famous Supervised Learning calculations, which is utilized for Classification as well as Regression . Be that as it may, principally, it is utilized for Classification problems in Machine Learning.
- The objective of the SVM calculation is to make the best line or choice limit that can isolate n-layered space into classes so we can undoubtedly put the new data of interest after in the right classification. This best choice limit is known as a hyper plane.

- SVM picks the outrageous focuses/vectors that assistance in making the hyperplane.

Navie Bayes

- Naive Bayes is a **probabilistic classifier** that returns the probability of a test point belonging to a class rather than the label of the test point.
- This algorithm is applicable for Classification tasks only, unlike many other ML algorithms which can typically perform Regression as well as Classification tasks.
- It uses conditional probability to calculate a product of individual probabilities of components.

IV. RESULTS

Fig. 2: Comparison of Accuracy Different ML Algorithms

The above figure represents the accuraciesof all the machine learning algorithms used in our present work.

Fig. 3: Crop Prediction

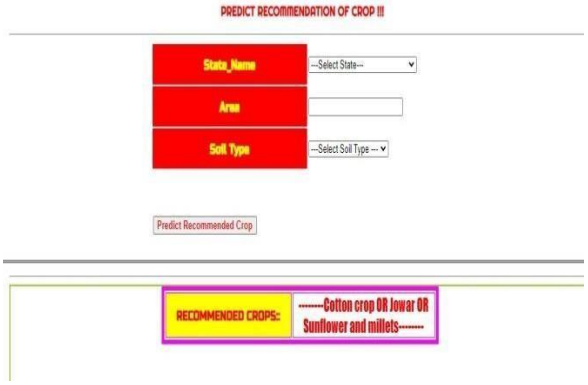

Fig. 4: Crop Recommendations

V. CONCLUSION& FUTURE SCOPE

At the point when we apply stacked relapse, the outcome has been so made do than when those models were applied independently. The result which has been displayed in figure is at present a web application, yet our future work would fabricate an application where the ranchers can utilize it as application and changing over the entire framework in their territorial languages.

REFERENCES

[1] “data.gov.in.” [Online]. Available: https://data.gov.in/

[2] Ananthara, M. G., Arunkumar, T., &Hemavathy, R. (2013, February). CRY—an improved crop yield prediction model using bee hive clustering approach for agricultural data sets. In 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (pp. 473-478).IEEE.

[3] Awan, A. M., & Sap, M. N. M. (2006, April). An intelligent system based on kernel methods for crop yield prediction. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 841-846).Springer, Berlin, Heidelberg.

[4] Bang, S., Bishnoi, R., Chauhan, A. S., Dixit, A. K., &Chawla, I. (2019, August). Fuzzy Logic based Crop Yield Prediction using Temperature and Rainfall parameters predicted through ARMA, SARIMA, and ARMAX models. In 2019 Twelfth International Conference on Contemporary Computing (IC3) (pp. 1-6). IEEE.

[5] Bhosale, S. V., Thombare, R. A., Dhemey, P. G., &Chaudhari, A. N. (2018, August). Crop Yield Prediction Using Data Analytics and Hybrid Approach. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (pp. 1-5).IEEE.

[6] Gandge, Y. (2017, December). A study on various data mining techniques for crop yield prediction.In 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT) (pp. 420-423).IEEE.

[7] Gandhi, N., Petkar, O., & Armstrong, L. J. (2016, July). Rice crop yield prediction using artificial neural networks.In 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR) (pp. 105-110).IEEE.

[8] Gandhi, N., Armstrong, L. J., Petkar, O., &Tripathy, A. K. (2016, July). Rice crop yield prediction in India using support vector machines.In 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 1-5).IEEE.

Author Profile:

N.KesavaRao, Associate Professor in the Departmentt of CSE at Narayana Engineering College,Gudur. He has published various patents and research articles.

Email: kesavarn@gmail.com

Desemsetty Ramya Sree a scholar who had successfully done an internship on Machine learning out of her own interest and was enthusiastic in learning new things apart from her regular academics.
Email: ramyadesemsetty@gmail.com

Chatla Mohan Rupa , final year student in the department of computer scienceand Enginnering at Narayana Engineering College, Gudur was an active volunteer for all the programmes and also a good learner.
Email: Mohanrupachatla19@gmail.com

Chava pavani , final year student in the department of computer science and Engineering at Narayana Engineering College, Gudur .She is a quick learner and also able to understand things easily.
Email: chpavaninaidu5@gmail.com

Cheethirala venkata sowdamini, final year student in the department of Computer Science and Engineering at Narayana Engineering College, Gudur has a good communication skills that are needed to get survive in any place. Email: chsow98@gmail.com