PREDICTING CROP YIELD RECOMMENDER SYSTEM USING MACHINE LEARNING TECHNICS

Mohammed Bilal
Department of IT
Nawab Shah Alam Khan College of Engineering and Technology
Hyderabad,India

Shahla Quraishi
Department of IT
Nawab Shah Alam Khan College of Engineering and Technology
Hyderabad,India

Zahoora Abid
Assistant Professor,Department of IT
Nawab Shah Alam Khan College of Engineering and Technology
Hyderabad,India

ABSTRACT — Horticulture and its unified areas are without a doubt the biggest suppliers of occupations in rustic India. The horticulture area is likewise a critical benefactor component to the nation's Gross Domestic Product (GDP). Gift to the nation is the mind-boggling size of the rural area. Notwithstanding, lamentable is the yield per hectare of harvests in contrast with worldwide norms. This is one of the potential reasons for a higher self-destruction rate among minimal ranchers in India. This paper proposes a feasible and easy to understand yield expectation framework for the ranchers. The proposed framework gives availability to ranchers by means of a portable application. GPS assists with distinguishing the client area. The client gives the region and soil type as info. AI calculations permit picking the most productive harvest list or anticipating the yield for a client chose crop. To foresee the harvest yield, chose Machine Learning calculations, for example, Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbor (KNN) are utilized. Among them, the Random Forest showed the best outcomes with 95% precision. Furthermore, the framework additionally recommends the best opportunity to utilize the composts to help up the yield.

I. INTRODUCTION

Horticulture has a broad history in India. As of late, India is positioned second in the homestead yield overall [15]. Farming related ventures, for example, ranger service and fisheries contributed for 16.6% of 2009 GDP and around half of the absolute labor force. Agribusiness' financial commitment to India's GDP is diminishing [1]. The harvest yield is the huge variable contributing in farming money related. The harvest yield relies upon numerous variables like climatic, geographic, natural, and monetary components [6]. It is hard for ranchers to conclude when and which yields to plant as a result of fluctuating business sector costs [7]. Refering to Wikipedia calculates India's self destruction rate goes from 1.4-1.8% per 100,000 populaces, throughout the course of recent years [15]. Ranchers know nothing about which yield to develop, and what is the perfect set-up to begin because of vulnerability in climatic circumstances. The utilization of different manures is likewise unsure because of changes in occasional climatic circumstances and essential resources like soil, water, and air. In this situation, the harvest yield rate is consistently declining [2]. The answer for the issue is to give a savvy easy to understand recommender framework to the ranchers. The harvest yield forecast is a critical issue in the horticulture area [3]. Each rancher attempts to realize crop yield and whether it measures up to their assumptions [4], consequently assessing the past experience of the rancher on the particular harvest foresee the yield [3]. Agribusiness yields depend basically on weather patterns, vermin, and readiness of collecting tasks. Precise data on crop history is basic for settling on choices on agribusiness risk the board [5]. In this paper, we have proposed a model that resolves these issues. The oddity of the proposed framework is to direct the ranchers to expand the harvest yield as well as recommend the most productive yield for the particular district. The proposed model gives crop determination in light of monetary and ecological circumstances, and benefit to expand the harvest yield that will therefore assist with fulfilling the rising need for the country's food supplies [8]. The proposed model predicts the harvest yield by concentrating on elements, for example, precipitation, temperature, region, season, soil type and so forth. The framework additionally assists with deciding the best chance to utilize manures. The current framework which suggests crop yield is either equipment based being expensive to keep up with, or not effectively available. The proposed framework recommends a portable based application that exactly predicts the most beneficial harvest by anticipating the yield. The utilization of GPS assists with recognizing the client area. The client gives a region under development and soil type as information sources. As per the necessity, the mode predicts the harvest yield for a particular yield. The model additionally suggests the most productive harvest and proposes the ideal opportunity to utilize the manures. The significant commitments of the paper are enrolled beneath, 1. Forecast of the harvest yield for explicit areas by executing different Machine Learning calculations, with a correlation of blunder rate and exactness. 2. An easy to understand versatile application to suggest the most productive yield. 3. A GPS based area identifier to recover the

precipitation assessment at the given region. 4. A recommender situation to propose the ideal opportunity for utilizing manures.

II. RELATED WORK

The means taken to support agribusiness basically includes imbuing mechanical mastery and creations to make the horticulture area more capable and rearranged for ranchers by anticipating the right yields utilizing all ML draws near. The paper examines different calculations like ANN, Fuzzy Network, and different information mining procedures with their benefits. Further test is to have every one of these consolidated constant datasets [9]. One of the early works fostered a devoted site to evaluate the effect of climate boundaries on crop creation in the distinguished locale of Madhya Pradesh [10]. The locale were chosen based on the district covered by the harvest. In light of these rules, the initial five top areas with a greatest yield region were picked. The premise of the harvests chose for the review was on winning yields in the chose regions. The harvests picked included maize, soybean, wheat and paddy, for which the yield for a constant time of 20 years of information, were classified. The exactness of the laid out model went from 76% to 90% for the picked crops with a typical precision of 82%. Another significant work really looks at the dirt quality and predicts the harvest yield alongside an appropriate suggestion of composts [11]. The Ph esteem and the area from the client were inputs utilized in this model. An API was utilized to anticipate the climate, temperature for the ongoing spot. The framework utilized both directed as well as solo ML calculations and analyzes the aftereffects of the two. A classifier that utilizes a ravenous technique to foresee the harvest yield was proposed in [12]. A choice tree classifier that utilizes a characteristic has been displayed to yield improved results. A troupe model proposed recommends incorporating the impacts of various models, which has been demonstrated to be commonly better compared to the singular models. Arbitrary timberlands gathering arrangement utilizes various choice tree models to foresee the harvest yield. The information are separated into two sets, like preparation information and test information, with a proportion of 67% and 33%, with which the mean and standard deviation are determined. This work likewise integrates the grouping of comparative yields to obtain the most dependable outcomes. Broad work has been done, and numerous ML calculations have been applied in the horticulture area. The greatest test in agribusiness is to increment ranch creation and proposition it to the end-client with the most ideal cost and quality. It is additionally seen that something like half of the homestead produce gets squandered, and it never arrives at the end-client. The proposed model recommends the techniques for limiting homestead produce wastage. One of the new works presents a model where the harvest yield is anticipated utilizing KNN calculations by making the bunches. It has been shown that KNN grouping demonstrated far superior to SVM or relapse [13]. In [17] predicts the harvest yield for the particular year with the assistance of cutting edge relapse strategies like Enet, Lasso and Kernel Ridge calculations. The Stacking relapse assisted with upgrading the precision of the calculations. The verifiable datasets are separated to recover the datasets for Maharashtra state utilizing Pandas profiling device. The harvest yield forecast model is planned utilizing multi-facet discernment brain organization and improved the exactness by changing inclination, weight and Adam analyzer. The proposed model purposes ANN with three-layer brain organization to foresee the harvest yield [18]. Managed learning approach is utilized to carry out crop yield expectation framework. Laid out the connection between's various qualities chose from the authentic which assists the framework with expanding the harvest yield [19]. Precipitation and temperature are two variables which impact the harvest yield. Intermittent Neural Network (RNN) and Long Short-Term Memory (LSTM) calculations applied on these time series information to improve the exactness [20]. ARMA (Auto Regressive Moving Average), SARIMA (Seasonal Auto Regressive Integrated Moving Average) and ARMAX (ARMA with exogenous factors) strategies are utilized to anticipate the temperature and precipitation utilizing verifiable information. The best model among them is utilized in the harvest yield forecast framework executed with fluffy rationale. Overcast cover and evapotranspiration are exogenous factors utilized in the proposed framework.

III. MODELS AND METHODOLOGY

In spite of numerous arrangements that have been as of late proposed, there are as yet open difficulties in making an easy-to-understand application regarding crop suggestion. The arrangement proposed here means to settle these restrictions, by fostering an easy-to-use application that considers the boundaries like precipitation, temperature, soil type and so forth that straightforwardly influence development. The primary goal is to acquire a superior assortment of yields that can be developed over the season. The proposed framework would assist with limiting the troubles looked by ranchers in picking a harvest and augment the yield in actuality to decrease the self-destruction rates [16]. The proposed model predicts the harvest yield for the informational indexes of the given area. Coordinating agribusiness and ML will add to

additional upgrades in the horticulture area by expanding the yields and advancing the assets in question. The information from earlier years are the vital components in anticipating current execution. Authentic information is gathered from different solid sources like data.gov.in, kaggle.com, and indianwaterportal.com. The informational indexes are gathered for Maharashtra and Karnataka areas. The information has different traits like state, region, year, season, sort of yield, a region under development, creation, and so on. The dirt sort is a characteristic in other datasets with state and locale determination. This dirt sort section is separated and converged into the fundamental informational collection. Likewise, temperature and normal precipitation are taken from a different dataset and added to the primary informational collections for the particular locale. The informational indexes are cleaned and pre-handled. The invalid qualities are supplanted with mean qualities. The downright qualities are changed over into marks prior to handling the calculations. The one hot encoding technique is utilized to manage unmitigated qualities in the informational collections. Figure 1 is the framework design of the proposed model. It's a versatile application that has two modules - the expectation module and the compost module. Portable Application offers various administrations. The rancher needs to enlist with the application through the enrollment interaction. When the enlistment is finished, the rancher can utilize the versatile application administrations. The expectation module predicts the harvest yield utilizing the chose credits from the informational collections for the particular yield. The foresee module additionally recommends the rancher with the best return crops. The compost module directs the rancher for the perfect opportunity to utilize the manure.

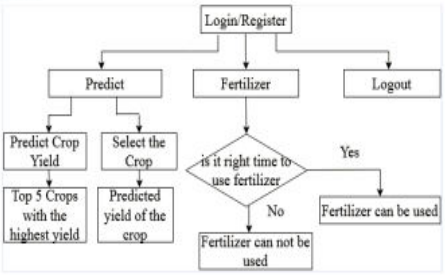

Fig. 1 System Architecture

Figure 2 illustrates the flow chart of the proposed system. It describes the whole process starting with the registration and various services provided by the mobile application.

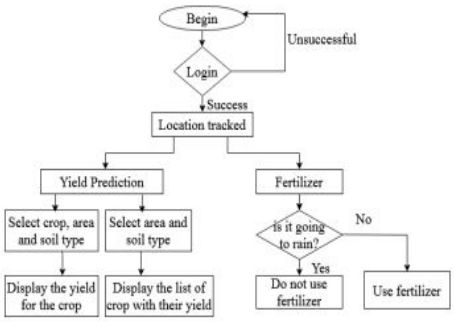

Fig. 2 Flow Chart

The very first step to use the services of the app is to register. During registration, the app locates the geographical location and identifies the region of the farmer using GPS. On successful login, the user can avail of two services. The first service is the yield prediction either for the selected crop or using a crop recommender system. The second service is the identification of the correct time to use the fertilizer. In the prediction service, the user needs to input the planned crop, soil type, and area under cultivation. The system predicts the yield for the specific crop selected. Figure 3 demonstrates the registration process to avail of the services of the app.

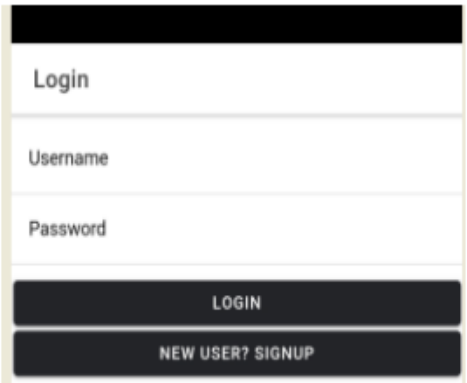

Fig. 3 Registration Process

If the farmer is not sure about the crop to be planned this year, he can use the crop recommender system. In the crop recommender system, the farmer must provide only soil type and area. The system lists the crops with their predicted yield. This makes farmers easy to decide on a crop to be planted. The timing of applying the fertilizer is very crucial. The farmer's effort and money will get wasted if the rain comes down too early. The proposed fertilizer usage service will guide the farmer on when to use the fertilizer. The model predicts the rain for the specific location for the next 14 days with Open Weather API. If the rainfall is more than 1.25 mm then it recommends as ‘not safe’ to use the fertilizers.

Figure 4 demonstrates the Block diagram of Experimental Implementation. The Graphical User Interface for the proposed model is developed with the Ionic Framework with JavaScript, Angular JS, and ReactJS. The system is built and deployed across multiple platforms such as iOS, Android, desktop, and the web as a Progressive Web Apps-all with one code base [14]. The datasets and resources required for the system are hosted on firebase. The machine learning approach is used for crop yield prediction. The patterns and correlations are discovered using ML approach. The model is trained using historical data sets where the past experience is used to represent the outcome. Various standard machine learning algorithms are used to predict yield. Among the selected algorithms, the Random Forest regression provided the best accuracy. Random Forest builds many decision trees and then blends them together to make the most accurate and stable predictions.

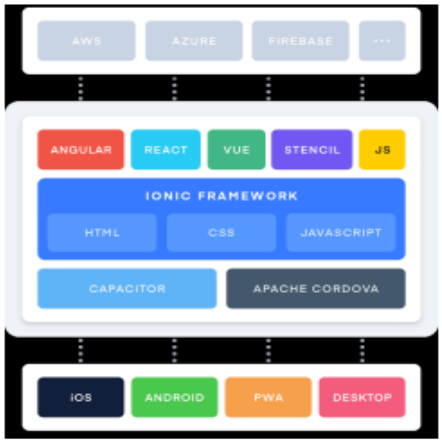

Fig. 4 Block diagram of Experimental Implementation

IV. RESULTS AND DISCUSSIONS

This part talks about the outcomes derived from chosen calculations for Maharashtra and Karnataka locales. The boundaries utilized for calculations are crop type, year, season, soil type, region, and area. For every one of the chose calculations, the exactness of the harvest yield forecast is analyzed. Arbitrary Forest calculation ended up being awesome for the given informational index with a precision of 95%. To anticipate the harvest yield, chose ML calculations like ANN, SVM, Multivariate Linear Regression, Random Forest, and KNN are utilized. Table1 shows the arranged consequences of the precision examination of different ML calculations. Figure 5 shows the graphical portrayal of the outcomes.

TABLE I : Accuracy vs Algorithm

Algorithm	Accuracy (%)
Artificial Neural Network (ANN)	86
Support Vector Machine (SVM)	75
Multivariate Linear Regression (MLR)	60
Random Forest (RF)	95
K Nearest Neighbor (KNN)	90

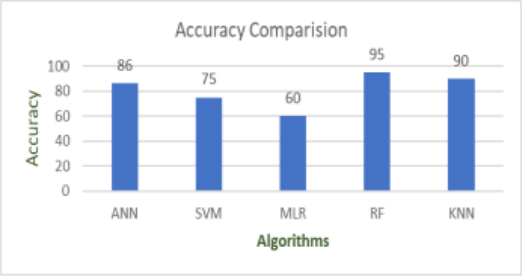

Fig. 5: Accuracy vs Algorithm

V. CONCLUSION

This paper featured the impediments of current frameworks and their functional utilization on yield forecast. Then strolls through a suitable yield forecast framework to the ranchers, a proposed framework gives availability to ranchers by means of a portable application. The versatile application incorporates numerous highlights that clients can use for the determination of a harvest. The inbuilt indicator framework assists the ranchers with foreseeing the yield of a given harvest. The inbuilt recommender framework permits a client investigation of the potential harvests and their respect take more instructed choices. For respect exactness, different AI calculations like Random Forest, ANN, SVM, MLR, and KNN were carried out and tried on the given datasets from the Maharashtra and Karnataka states. The different calculations are contrasted and their exactness. The outcomes acquired demonstrate that Random Forest Regression is awesome among the arrangement of standard calculations utilized on the given datasets with an exactness of 95%. The proposed model additionally investigated the planning of applying composts and suggests suitable length.

VI. FUTURE ENHANCEMENT

The future work will be centered around refreshing the datasets every once in a while, to deliver precise forecasts, and the cycles can be robotized. One more usefulness to be executed is to give the right kind of manure for the given yield and area. To carry out this careful investigation of accessible composts and their relationship with soil and environment should be

finished. An examination of accessible measurable information should be finished.

REFERENCES

[1] Umamaheswari S, Sreeram S, Kritika N, Prasanth DJ, "BIoT: Blockchain-based IoT for Agriculture", 11th International Conference on Advanced Computing (ICoAC), 2019 Dec 18 (pp. 324-327). IEEE.

[2] Jain A. "Analysis of growth and instability in the area, production, yield, and price of rice in India", Journal of Social Change and Development, 2018;2:46-66

[3] Manjula E, Djodiltachoumy S, "A model for prediction of crop yield" International Journal of Computational Intelligence and Informatics, 2017 Mar;6(4):2349-6363.

[4] Sagar BM, Cauvery NK., "Agriculture Data Analytics in Crop Yield Estimation: A Critical Review", Indonesian Journal of Electrical Engineering and Computer Science, 2018 Dec;12(3):1087-93.

[5] Wolfert S, Ge L, Verdouw C, Bogaardt MJ, "Big data in smart farming– a review. Agricultural Systems", 2017 May 1;153:69-80.

[6] Jones JW, Antle JM, Basso B, Boote KJ, Conant RT, Foster I, Godfray HC, Herrero M, Howitt RE, Janssen S, Keating BA, "Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. Agricultural systems", 2017 Jul 1;155:269-88.

[7] Johnson LK, Bloom JD, Dunning RD, Gunter CC, Boyette MD, Creamer NG, "Farmer harvest decisions and vegetable loss in primary production. Agricultural Systems", 2019 Nov 1;176:102672. [8] Kumar R, Singh MP, Kumar P, Singh JP, "Crop Selection Method to maximize crop yield rate using a machine learning technique", International conference on smart technologies and management for computing, communication, controls, energy, and materials (ICSTM), 2015 May 6 (pp. 138-145). IEEE.

[9] Sriram Rakshith.K, Dr.Deepak.G, Rajesh M, Sudharshan K S, Vasanth S, Harish Kumar N, "A Survey on Crop Prediction using Machine Learning Approach", In International Journal for Research in Applied Science & Engineering Technology (IJRASET), April 2019, pp(3231- 3234)

[10] Veenadhari S, Misra B, Singh CD, "Machine learning approach for forecasting crop yield based on climatic parameters", In 2014 International Conference on Computer Communication and Informatics, 2014 Jan 3 (pp. 1-5). IEEE.

[11] Ghadge R, Kulkarni J, More P, Nene S, Priya RL, "Prediction of crop yield using machine learning", Int. Res. J. Eng. Technol. (IRJET), 2018 Feb;5.

[12] Priya P, Muthaiah U, Balamurugan M, "Predicting yield of the crop using machine learning algorithm", International Journal of Engineering Sciences & Research Technology, 2018 Apr;7(1):1-7. [13] S. Pavani, Augusta Sophy Beulet P., "Heuristic Prediction of Crop Yield Using Machine Learning Technique", International Journal of Engineering and Advanced Technology (IJEAT), December 2019, pp (135-138)

[14] https://web.dev/progressive-web-apps/

[15] https://www.wikipedia.org/

[16] Plewis I, "Analyzing Indian farmer suicide rates", Proceedings of the National Academy of Sciences, 2018 Jan 9;115(2): E117.

[17] Nishant, Potnuru Sai, Pinapa Sai Venkat, Bollu Lakshmi Avinash, and B. Jabber. "Crop Yield Prediction based on Indian Agriculture using Machine Learning." In 2020 International Conference for Emerging Technology (INCET), pp. 1-4. IEEE, 2020.

[18] Kale, Shivani S., and Preeti S. Patil. "A Machine Learning Approach to Predict Crop Yield and Success Rate." In 2019 IEEE Pune Section International Conference (PuneCon), pp. 1-5. IEEE, 2019. [19] Kumar, Y. Jeevan Nagendra, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi. "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, 2020.

[20] Nigam, Aruvansh, Saksham Garg, Archit Agrawal, and Parul Agrawal. "Crop yield prediction using machine learning algorithms." In 2019 Fifth International Conference on Image Information Processing (ICIIP), pp. 125-130. IEEE, 2019.

[21] Bang, Shivam, Rajat Bishnoi, Ankit Singh Chauhan, Akshay Kumar Dixit, and Indu Chawla. "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, 2019.