MagiCrop: Smart Crop and Fertilizer Recommendation System using ML Algorithms

[1]Rahul R. Mahurkar, [2]Pankaj R. Chopade, [3]Omkar A. Kudile, [4]Santosh Sharma, [5]Sharad Adsure

[1,2,3,4]Software Engineer, [5]Professor Comp. department
[1,2,3,4,5]Computer Department, Pune, India

Abstract: Farming is a significant supporter of the Indian economy. The standard Indian populace depends either verifiably unequivocally on agribusiness for their occupation. In this manner, it assumes an imperative part of the country. Avast larger part of the farmers has faith in relying upon their instinct to choose which harvest to plant in the specific season and what manures to utilize. This has been going on since old occasions. Thus, a solitary farmer can't be relied upon to consider every one of the countless variables that add to crop development before arriving at an agreement about which one to develop. A solitary misinformed choice by the farmer can have unfortunate implications on both himself and the country's agrarian economy. In this paper, we present an insightful framework called MagiCrop, which expects to help the farmer in settling on educated choices about the harvests to develop and the manures to use to boost the yield, depending upon natural factors and contents of the soil in that particular region.

Index Terms - **Crop prediction, Fertilizer recommendation, Smart farming, Machine learning.**

I. INTRODUCTION

Maharashtra went through a few vacillations last year as for the retail cost of onions. The cost expanded from Rs. 26 for each kilo in the primary portion of the year to an incredible Rs. 50 for every kilo in August. Noticing the shoot in the value, a considerable lot of the ranchers in the state chose to develop onions on their homestead, in the expectation of making excessive benefits. While this brought about plentiful stock in specific areas of Maharashtra, numerous different locales endured bombed crop yield because of troublesome conditions for developing onions. A resulting deficiency again before long had cruel repercussions on the existences of everyday person, as middle-class families could presently don't manage the cost of onion-an as often as possible utilized ware in their kitchen.

This model simply proceeds to show that a rancher's choice about which yield to develop is by and large obfuscated by his instinct and other unimportant elements like making moment benefits, absence of mindfulness about market interest, overestimating a dirt's capability to help a specific harvest, etc. An extremely misinformed choice concerning the rancher could put a huge strain on his family's monetary condition. Maybe this could be one of the numerous reasons adding to the endless self-destruction instances of ranchers that we hear from media consistently. In a nation like India, where horticulture and related area adds to around 20.4 per cent of its Gross Value Added (GVA), a particularly mistaken judgment would have negative ramifications on the rancher's family, however the whole economy of a district. Consequently, we have distinguished a rancher's predicament about which harvest to develop during a specific season, as a grave one.

The need of great importance is to plan a framework that could give prescient bits of knowledge to the Indian ranchers, subsequently assisting them with settling on an educated choice about which yield to develop. In light of this, we propose a Magi Crop -a clever framework that would think about ecological boundaries (temperature, precipitation, homestead's scope, longitude, elevation and distance from the ocean) and soil attributes (pH esteem, soil type and thickness of spring and dirt) before prescribing the most reasonable harvest to the client. This model would take contribution from another suggestion framework, called Rainfall Predictor, which would foresee the month-wise precipitation of the following year for the specific client's area.

II. LITERATURE REVIEW

An ever-increasing number of scientists have started to recognize this issue in Indian farming and are progressively devoting their time and endeavors to help ease the issue. Authors[3] utilize Regularized Greedy Forest to decide an suitable harvest arrangement at a given time stamp. The creators of [4] have proposed a model that utilizes verifiable records of meteorological information as preparing set. Model is prepared to distinguish climate conditions that are obstacle for the creation of apples. It then, at that point productively predicts the yield of apples based on month-to-month climate designs. The impact of temperature on the sugar substance of apples is likewise taken into record to distinguish likely measure of harmed yield.

The utilization of a few algorithms like Artificial Neural Organization, K-Nearest Neighbors, and Regularized Greedy Forest is exhibited in [5] to choose a harvest dependent on the forecast yield rate, which, is affected by various boundaries. Extra highlights remembered for the framework are pesticide forecast and web-based exchanging dependent on farming wares. Another smart model, introduced in [6], takes into account the forecast of soil credits like phosphorous substance. Here, the creators utilize, diverse arrangement methods like Credulous Bayes, C4.5, Linear Regression and Least Median Square to accomplish high forecast exactness. This framework can be extremely valuable for ranchers to decide the reasonableness of the soil to help a specific yield.
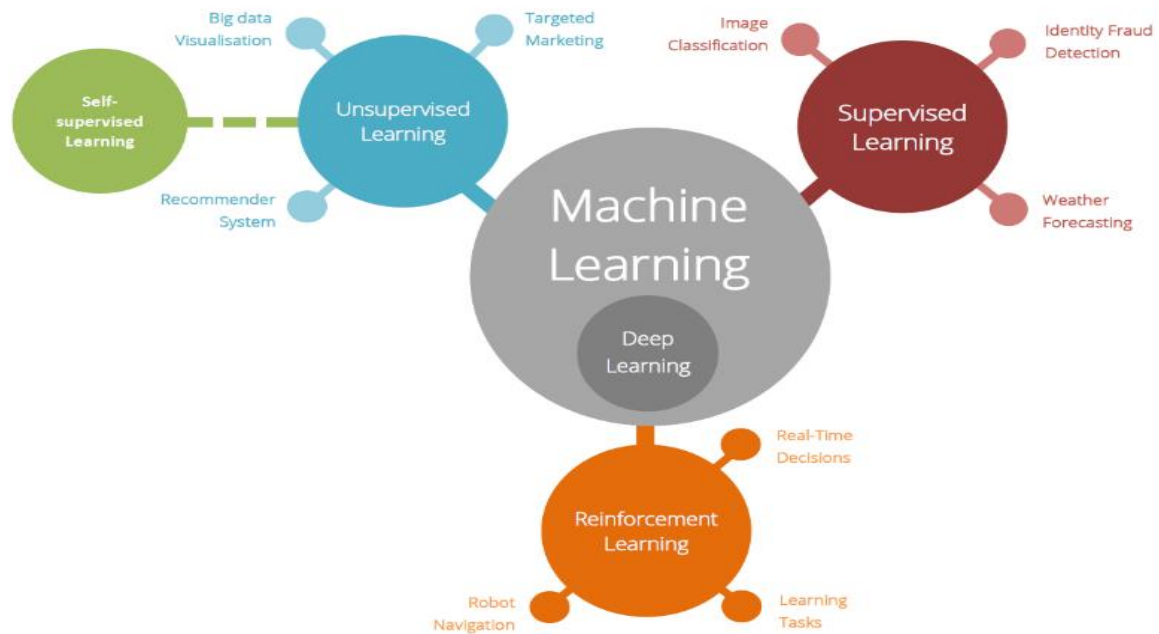
III. METHODOLOGY

Fig.01 Machine Learning

3.1 MACHINE LEARNING:

Machine learning (ML) may be a kind of AI (AI) that permits package applications to become additional correct at predicting outcomes while not being expressly programmed to try and do thus. Machine learning algorithms use historical knowledge as input to predict new output values.

Recommendation engines are a typical use case for machine learning. different widespread uses embody fraud detection, spam filtering, malware threat detection, business method automation (BPA) and prognostic maintenance. Machine learning is very important as a result of it provides enterprises a read of trends in client behavior and business operational patterns, further as supports the event of recent merchandise. several of today's leading corporations, like Facebook, Google and Uber, build machine learning a central a part of their operations. Machine learning has become a major competitive human for several corporations.

SUPERVISED LEARNING:

Supervised learning: during this sort of machine learning, information scientists provide rules with labelled coaching information and outline the variables they require the algorithm to assess for correlations. each the input and also the output of the rule is nominative. Supervised machine learning needs the info person to coach the rule with each labelled inputs and desired outputs. supervised learning algorithms area unit sensible for the subsequent tasks: Binary classification: Dividing information into 2 classes. Multi-class classification: selecting between quite 2 varieties of answers. Regression modelling: Predicting continuous values. Ensembling : Combining the predictions of multiple machine learning models to supply Associate in Nursing correct prediction.

UNSUPERVISED LEARNING:

Unsupervised learning: this kind of machine learning involves algorithms that train on unlabeled information. The rule scans through information sets trying to find any purposeful affiliation. the information that algorithms train on still because the predictions or recommendations they output square measure planned. Unsupervised machine learning algorithms don't need information to be labelled. They sift through unlabeled information to seem for patterns that may be wont to cluster information points into subsets. Most styles of deep learning, together with neural networks, square measure unsupervised algorithms. unsupervised learning algorithms square measure smart for the subsequent tasks:

Clustering: rending the dataset into teams supported similarity. Anomaly detection: distinctive uncommon information points in an exceedingly information set. Association mining: distinctive sets of things in an exceedingly information set that regularly occur along. Dimensionality reduction: Reducing the amount of variables in an exceedingly information set.

REINFORCEMENT LEARNING:

Reinforcement learning is an area of Machine Learning. It is about taking suitable action to maximize reward in a particular situation. It is employed by various software and machines to find the best possible behavior or path it should take in a specific situation. Reinforcement learning differs from the supervised learning in a way that in supervised learning the training data has the answer key with it so the model is trained with the correct answer itself whereas in reinforcement learning, there is no answer but the reinforcement agent decides what to do to perform the given task. In the absence of a training dataset, it is bound to learn from its experience. The existing system of agriculture is mostly based on trial-and-error cultivation of the farmers based on their experiences, up until the date. Also, not many farmers are connected to technology. Thus, this issue has been only found in the number of review papers but no system has existed for that specific purpose. There are also some drawbacks of this existing system. Such as, it takes a lot of time to exactly know what kind of crop to grow, High uncertainty regarding crop yield, and it is also costs a lot to hire an expert for this purpose.

IV. PROPOSED APPROACH

Our proposed approach is mainly based on wiping out the previously mentioned downsides, we propose our framework, MagiCrop which contemplates every one of the boundaries including soil contents, precipitation and area to anticipate crop development. Following diagram outlines the framework's engineering.
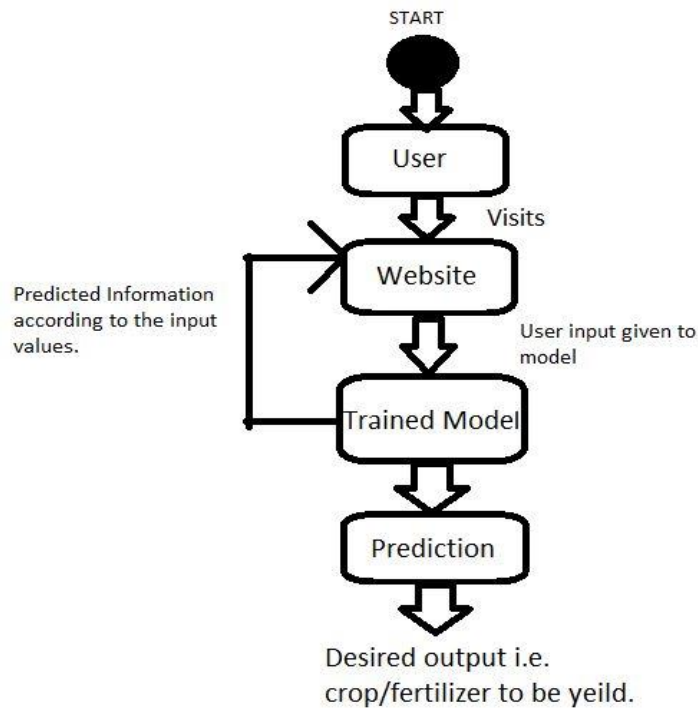

Fig.02 Proposed System

As illustrated in the given diagram of the framework's architecture, a user who might be anyone from a newly started farmer or even an experienced one, visits the website, they are welcomed with an intuitive and user-friendly interface. This is where, one needs to fill basic information asked by the system, which can be easily obtained from the soil test done at the land where crops are to be grown. Also, some additional information like rainfall and location needs to be entered along with the nutrients present in the soil. And once that is done, main working of the system begins. The input entered in the website goes to the trained model, where the data is checked and compared with the predefined functions programmed by the programmers. This helps to predict the required information according to the input values, which are present in the dataset. Bigger the size of the dataset, more accurate output can be provided based on that. And after prediction of the most suitable crop to grow in these conditions, it is displayed on the screen. Similar, process takes place while one considers what fertilizers are better to use for their crops. And thus, one can also gain insight depending on his or her inputs.

4.1 MACHINE LEARNING ALGORITHMS:

In the proposed model of our system, a number of algorithms have been used according to their accuracy. Following are the algorithms that have been used in the system: Random Forest (R.F), Decision Tree (DT) and Naive Bayes Theorem. (N.B.T)

DECISION TREE:

It's anything but a managed learning calculation where credits and class names are addressed utilizing a tree. Here, root credits are contrasted and the record's characteristic and, in this way, contingent on the correlation, another hub is reached. This correlation is proceeded until a leaf hub with an anticipated class esteem is reached. Thusly, a demonstrated choice tree is extremely effective for forecast purposes. Decision trees area unit one among the most effective sorts of learning algorithms supported varied learning strategies. They boost prognosticative models with accuracy, ease in interpretation, and stability. The tools are effective in fitting non-linear relationships since they're capable of finding data-fitting challenges, like regression and classifications. One of the applications of call trees involves evaluating prospective growth opportunities for businesses supported historical knowledge. Historical knowledge on sales is employed in call trees that will cause creating radical changes within the strategy of a business to assist aid growth and growth.
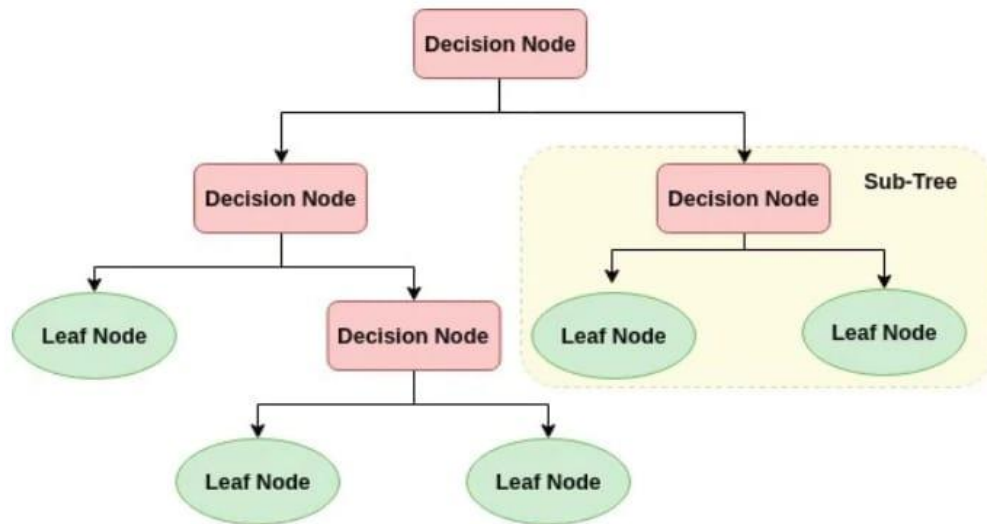
Fig.03 Decision Tree

RANDOM FOREST:

It is AN ensemble technique of learning that's ordinarily used for each classification and regression. so as to coach the model to perform prediction mistreatment this algorithmic program, the test features should be older the foundations of every randomly created tree. As a result of this, a special target is going to be predicted by every random forest for identical check feature. Then, votes area unit calculated on the idea of every expected target. the ultimate prediction of the rule is that the highest votes expected target. the actual fact that random forest rule can with efficiency handle missing values which the classifier can ne'er over-fit the model area unit vast edges for victimization this algorithm.

Random Forest is a mainstream AI calculation that has a place with the managed learning method. It tends to be utilized for both Classification and Regression issues in ML. It depends on the idea of gathering realizing, which is a cycle of joining numerous classifiers to take care of an unpredictable issue and to improve the presentation of the model. The more noteworthy number of trees in the woods prompts higher precision and forestalls the issue of overfitting. Since the irregular woodland joins various trees to anticipate the class of the dataset, it is conceivable that some choice trees may foresee the right yield, while others may not. In any case, together, every one of the trees anticipate the right yield. Consequently, beneath are two presumptions for a superior Random woodland classifier: There ought to be some genuine qualities in the component variable of the dataset so the classifier can foresee exact outcomes instead of a speculated result. The expectations from each tree should have exceptionally low relationships.
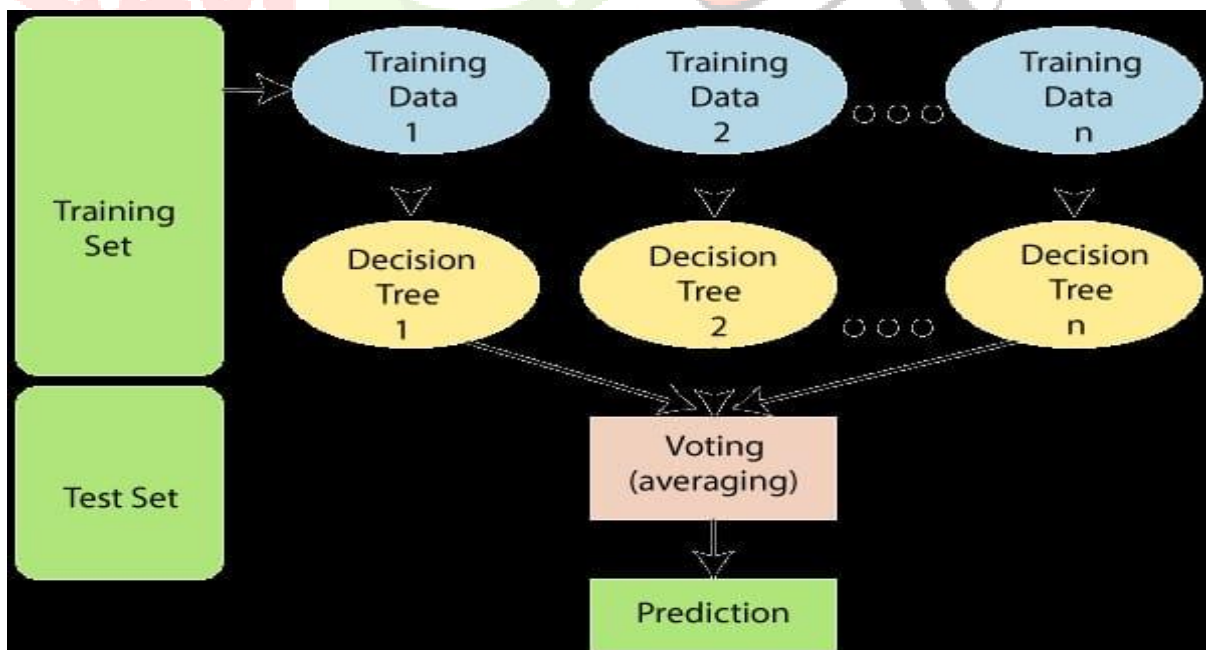

Fig.04 Random Forest

NAIVE BAYES THEORM:

Naive Thomas Bayes Classifier is one amongst the straightforward and simplest Classification algorithms that helps in building the quick machine learning models which will create fast predictions. it's a probabilistic classifier, which suggests it predicts on the premise of the likelihood of associate object. is appropriate for resolution multi-class prediction issues. If its assumption of the independence of options holds true, it will perform higher than alternative models and needs abundant less coaching information. Naive Bayes is healthier suited to categorical input variables than numerical variables. Naive Classifier could be an easy classification model that assumes very little to nada concerning the matter and also the performance of that provides a baseline by that all alternative models evaluated on a dataset is compared. There are totally different methods which will be used for a naive classifier, and a few are higher than others, betting on the dataset and also the selection of performance measures. the foremost common performance live is classification accuracy and customary naive classification methods, as well as random dead reckoning category labels, randomly selecting labels from a coaching dataset, and employing a majority category label.

It is valuable to foster a little likelihood structure to ascertain the normal presentation of a given innocent characterization methodology and to perform investigations to affirm the hypothetical assumptions. These activities give an instinct both to the conduct of gullible order calculations as a rule, and the significance of setting up an exhibition pattern for a grouping task.

CONCLUSION:

In this paper, we have effectively proposed and executed a smart harvest recommendation framework, which can be handily utilized by farmers all over world. This framework would help the farmers in settling on an educated choice about which harvest to develop contingent upon an assortment of ecological and geological components. We have also implemented another function of fertilizer recommendation that is suitable for the crops grown to gain a great yield based on the geographical location of the region. The high accuracies given by both these models make them extremely productive for all useful and continuous purpose**s.**

The model proposed in this paper can be additionally stretched out later on to fuse an element to anticipate crop infections. This would guarantee boosted yield as the choice about which yield to develop would now likewise rely on which yield was gathered in the past cycle. Besides, crop supply and demand just as other monetary markers like homestead reap costs and retail costs can likewise be considered as boundaries to the Crop Suitability Predictor model. This would give an all-encompassing expectation not just based on natural and geological elements, yet in addition relying upon the monetary angles.

REFERENCES:

[1] J he, "Bayesian dual network for recommendation", vol (13)

[2] "A comprehensive survey and analysis of existing evaluation", vol 24, Pedro'c. Camps

[3] "Leaflet-a JavaScript library for interactive maps", S. Liu, mar18

[4] "Prediction of crop yield using MLX", International research journal of egg, vol5

[5] "Application of data analytics in agriculture", V. R. Malovic, 2014

[6] "Measurement of NPK from pH value, 2016, international research journal of engineering and technology

[7] " Development of novel bio inspired framework ", Dubai, Y C Liu, 2018

[8] Komal Abhang," Soil analysis and crop fertility prediction", 2018