



Crop Yield Prediction Based on Indian Agriculture Using Machine Learning

Medam Swarnalatha¹, K.Sasikanth²

¹M.Tech, Student, Chalapathi Institute of Engineering and Technology, Lam, Andrapradesh, India

²Assistant Professor, Chalapathi Institute of Engineering and Technology, Lam, Andrapradesh, India

ABSTRACT

Agriculture is one of the significant incomes creating areas of India and a wellspring of endurance. Various occasional, monetary and organic examples impact the harvest creation however capricious changes in these examples lead to an extraordinary misfortune to ranchers. These dangers can be diminished when reasonable methodologies are utilized on information identified with soil type, temperature, air pressing factor, mugginess and yield type. While, harvest and climate anticipating can be anticipated by getting valuable experiences from these agrarian information that guides ranchers to settle on the yield they might want to plant for the impending year prompting greatest benefit. This paper presents an overview on the different calculations utilized for climate, crop yield, and harvest cost expectation.

Keywords: Agriculture, Random forest algorithm, Machine Learning, Crop, Weather.

1.INTRODUCTION:

The horticulture area needs a colossal up degree to endure the changing states of the Indian economy. Alongside the advances in machines and advances utilized in cultivating, helpful and precise data about various matters additionally assumes a critical part in it. This data is being assembled by the utilization of distant sensors, satellite pictures, studies and so forth This data alongside the information on subject specialists and scientists ought to be promptly accessible to the ranchers to misuse its possible worth. This situation basically focuses on climate gauging, crop yield expectation and harvest cost estimating. These variables assist the ranchers with developing the best food yields and raise the correct animals with understanding of natural segments. Additionally, the ranchers can adjust to environment changes somewhat by moving planting dates, picking assortments with various development length, or changing yield turns. For exploratory examination, the factual numeric information identified with agribusiness is embraced.

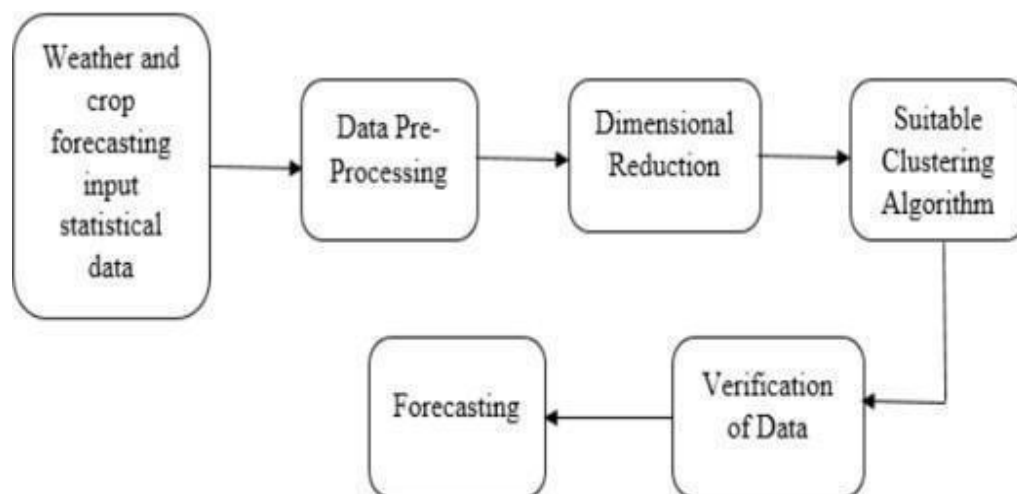


Figure 1. General Forecasting block diagram

Though, the bunching based methods and directed calculations are used for dealing with the gathered factual information. Furthermore, the appropriate order strategies like Help Vector Machine (SVM), neural organizations are utilized for better arrangement results. India's agribusiness comprises various yields, with the significant harvests of rice and wheat. Indian ranchers develop heartbeats, sugarcane and furthermore, non-food things like cotton, tea, espresso, etc. These methods will help in foreseeing the precipitation, crop yield determining and cost forecast of harvests as demonstrated in Figure 1.

2 METHODS:

2.1 Dataset and Pre-processing:

We required a dataset of phony and real profiles. Different ascribes remembered for dataset are number of companions, devotees, status tally. Dataset is separated into preparing and testing information. Arrangement calculations are prepared utilizing preparing dataset and testing dataset is utilized to decide effectiveness of calculation. From the dataset utilized, 80% of the two profiles (authentic and counterfeit) are utilized to set up a preparation dataset and 20% of the two profiles are utilized to set up a testing dataset.

2.2 Feature Selection:

Highlights are chosen to apply arrangement calculations. The arrangement calculation is examined further. Qualities are chosen as highlights in the event that they are not reliant on different ascribes and they increment effectiveness of the characterization. The highlights that we have picked are talked about further.

After choice of characteristics, the dataset of profiles that are as of now delegated phony or authentic are required for the preparation motivation behind the characterization calculation. We have utilized an openly accessible dataset of 1337 phony clients and 1481 authentic clients comprising different ascribes including name, status check, number of companions, supporters tally, top picks, dialects known and so forth.

2.3 Classification

Characterization is the way toward arranging an information object into classifications called classes dependent on highlights/credits related with that information object. Grouping utilizes a classifier, a calculation that measures the credits of every information item and yields a class dependent on this data. In this task, we use the Support Vector Machine as a classifier. Backing Vector Machine is an exquisite and strong strategy for characterization on an enormous informational collection much the same as the informational indexes of Social Network with a few great many profiles. The utilization case graph for crop yield expectation appears in Figure 2.

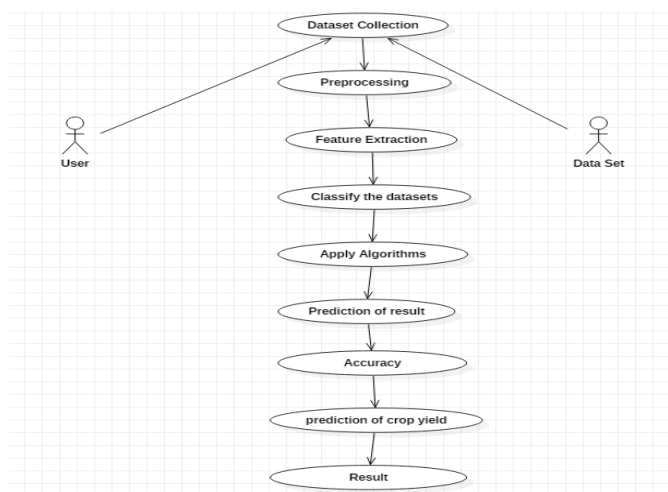


Figure 2. Use case diagram for crop yield prediction

2.4 Random Forest:

Random forest is a type of supervised machine learning algorithm based on ensemble learning. Ensemble learning is a type of learning where you join different types of algorithms or same algorithm multiple times to form a more powerful prediction model. The random forest algorithm combines multiple algorithms of the same type i.e., multiple decision trees, resulting in a forest of trees, hence the name "Random Forest". The random forest algorithm can be used for both regression and classification tasks.

2.5 Working of Random Forest

Coming up next are the essential advances engaged with playing out the arbitrary woodland calculation

1. Pick N arbitrary records from the dataset.

2. Build a choice tree dependent on these N records.
3. Choose the quantity of trees you need in your calculation and rehash stages 1 and 2.
4. For arrangement issues, each tree in the woodland predicts the classification to which the new record has a place. At last, the new record is allocated to the class that successes the lion's share vote.

Preferences of utilizing irregular woodland: geniuses of utilizing arbitrary timberland for grouping and relapse.

1. The irregular timberland calculation isn't one-sided, since, there are various trees and each tree is prepared on a subset of information. Essentially, the arbitrary timberland calculation depends on the force of "the group"; in this way, the generally speaking biasedness of the calculation is decreased.
2. This calculation is truly steady. Regardless of whether another information point is presented in the dataset the general calculation isn't influenced much since new information may affect one tree, yet it is difficult for it to affect all the trees.
3. The irregular timberland calculation functions admirably when you have both clear cut and mathematical highlights.
4. The arbitrary backwoods calculation likewise functions admirably when information has missing qualities or it has not been scaled well.

Three normal information pre-preparing steps are :

- **Formatting:** The information you have chosen may not be in an arrangement that is reasonable for you to work with. The information might be in a social data set and you might want it in a level record, or the information might be in a restrictive document arrangement and you might want it in a social data set or a content document.
- **Cleaning:** Cleaning information is the expulsion or fixing of missing information. There might be information cases that are fragmented and don't convey the information you trust you need to address the issue. These cases may be taken out. Furthermore, there might be touchy data in a portion of the credits and these ascribes may should be anonymized or eliminated from the information.
- **Sampling:** There might be definitely more chosen information accessible than you need to work with. More information can bring about any longer running occasions for calculations and bigger computational and memory necessities. You can take a more modest delegate test of the chosen information that might be a lot quicker for investigating and prototyping arrangements prior to considering the entire dataset.

Machine Learning:

AI is the mind where all the learning happens. The manner in which the machine learns is like the individual. People gain as a matter of fact. The more we know, the more effectively we can anticipate. By relationship, when we face an obscure circumstance, the probability of achievement is lower than the known circumstance. Machines are prepared for the equivalent. To make a precise expectation, the machine sees a model. At the point when we give the machine a comparative model, it can sort out the result. Notwithstanding, similar to a human, if it feeds a formerly inconspicuous model, the machine experiences issues to anticipate. The center goal of AI is learning and induction. Above all else, the machine learns through the disclosure of examples. This disclosure is made with gratitude to the information. One vital piece of the information researcher is to pick cautiously which information to give to the machine. The rundown of characteristics used to tackle an issue is known as a component vector. You can consider an element vector a subset of information that is utilized to handle an issue.

The machine utilizes some extravagant calculations to improve on the truth and change this revelation into a model. In this manner, the learning stage is utilized to depict the information and sum up it into a model.

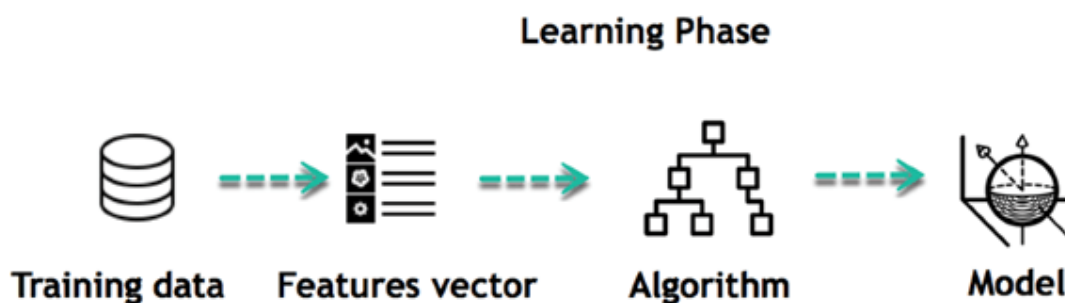


Figure 3. Learning Phase

For instance, the machine is trying to understand the relationship between the wage of an individual and the likelihood to go to a fancy restaurant. It turns out the machine finds a positive relationship between wage and going to a high-end restaurant: This is the model

When the model is built, it is possible to test how powerful it is on never-seen-before data. The new data are transformed into a features vector, go through the model and give a prediction. This is all the beautiful part of machine learning. There is no need to update the rules or train again the model. You can use the model previously trained to make inference on new data.

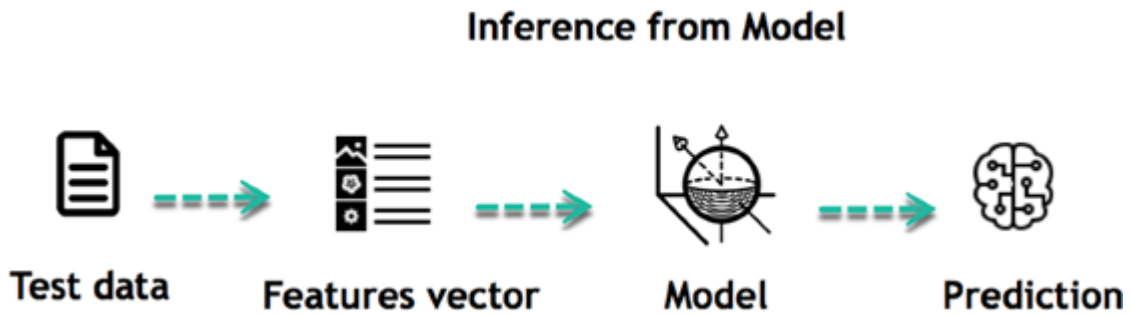


Figure 4. Inference from Model

The life of Machine Learning programs is straightforward and can be summarized in the following points:

1. Define a question
2. Collect data
3. Visualize data
4. Train algorithm
5. Test the Algorithm
6. Collect feedback
7. Refine the algorithm
8. Loop 4-7 until the results are satisfying
9. Use the model to make a prediction

When the calculation gets the hang of reaching the correct inferences, it applies that information to new arrangements of information. AI can be gathered into two expansive learning assignments: Supervised and Unsupervised. There are numerous different calculations Directed Learning: A calculation utilizes preparing information and criticism from people to get familiar with the relationship of offered contributions to a given yield. For example, an expert can utilize showcasing cost and climate conjecture as information to anticipate the deals of jars.

You can utilize administration when the yield information is known. The calculation will anticipate new information. There are two classifications of managed learning:

- Classification task
- Regression task

Grouping: Imagine you need to foresee the sexual orientation of a client for a business. You will begin gathering information on the stature, weight, work, compensation, buying bushel, and so forth from your client data set. You know the sexual orientation of every one of your clients, it must be male or female. The target of the classifier will be to allot a likelihood of being a male or a female (i.e., the mark) in light of the data (i.e., highlights you have gathered). At the point when the model figured out how to perceive male or female, you can utilize new information to make an expectation. For example, you just got new data from an obscure client, and you need to know whether it is a male or female. In the event that the classifier predicts male = 70%, it implies the calculation makes certain at 70% that this client is a male, and 30% it is a female.

The mark can be of at least two classes. The above model has just two classes, however on the off chance that a classifier needs to foresee object, it has many classes (e.g., glass, table, shoes, and so on each article addresses a class)

Relapse: When the yield is a persistent worth, the errand is a relapse. For example, a monetary investigator may have to conjecture the estimation of a stock dependent on a scope of highlight like value, past stock exhibitions, macroeconomics record. The framework will be prepared to assess the cost of the stocks with the most minimal conceivable blunder.

Algorithm Name	Description	Type
Linear regression	Finds a way to correlate each feature to the output to help predict future values.	Regression

Logistic regression	Extension of linear regression that's used for classification tasks. The output variable is binary (e.g., only black or white) rather than continuous (e.g., an infinite list of potential colours)	Classification
Decision tree	Highly interpretable classification or regression model that splits data-feature values into branches at decision nodes (e.g., if a feature is a colour, each possible colour becomes a new branch) until a final decision output is made	Regression Classification
Naive Bayes	The Bayesian method is a classification method that makes use of the Bayesian theorem. The theorem updates the prior knowledge of an event with the independent probability of each feature that can affect the event.	Regression Classification
Support vector machine	Support Vector Machine, or SVM, is typically used for the classification task. SVM algorithm finds a hyperplane that optimally divided the classes. It is best used with a non-linear solver.	Regression Classification
Random forest	The algorithm is built upon a decision tree to improve the accuracy drastically. Random forest generates many times simple decision trees and uses the 'majority vote' method to decide on which label to return. For the classification task, the final prediction will be the one with the most vote; while for the regression task, the average prediction of all the trees is the final prediction.	Regression Classification
AdaBoost	Classification or regression technique that uses a multitude of models to come up with a decision but weighs them based on their accuracy in predicting the outcome	Regression Classification
Gradient-boosting trees	Gradient-boosting trees is a state-of-the-art classification/regression technique. It is focusing on the error committed by the previous trees and tries to correct it.	Regression Classification

Unsupervised Learning: In unsupervised learning, an algorithm explores input data without being given an explicit output variable (e.g., explores customer demographic data to identify patterns) You can use it when you do not know how to classify the data, and you want the algorithm to find patterns and classify the data for you

Algorithm	Description	Type
K-means clustering	Puts data into some groups (k) that each contains data with similar characteristics (as determined by the model, not in advance by humans)	Clustering
Gaussian mixture model	A generalization of k-means clustering that provides more flexibility in the size and shape of groups (clusters)	Clustering
Hierarchical clustering	Splits clusters along a hierarchical tree to form a classification system. Can be used for Cluster loyalty-card customer	Clustering
Recommender system	Help to define the relevant data for making a recommendation.	Clustering
PCA/T-SNE	Mostly used to decrease the dimensionality of the data. The algorithms reduce the number of features to 3 or 4 vectors with the highest variances.	Dimension Reduction

3. RESULTS AND DISCUSSIONS:

AI, you need less information to prepare the calculation than profound learning. Profound learning requires a broad and assorted arrangement of information to recognize the fundamental design. Additionally, AI gives a quicker prepared model. Most exceptionally profound learning design can require days to seven days to prepare. The benefit of profound learning over AI is it is exceptionally exact. You don't have to comprehend what

highlights are the best portrayal of the information; the neural organization figured out how to choose basic highlights. In AI, you need to decide for yourself what highlights to remember for the model.

Tensor Flow:

The most renowned profound learning library on the planet is Google's TensorFlow. Google item utilizes AI taking all things together with its items to improve the web crawler, interpretation, picture inscribing or proposals.

To give a solid model, Google clients can encounter a quicker and more refined hunt with AI. On the off chance that the client types a watchword in the inquiry bar, Google gives a proposal about what could be the following word.

Google needs to utilize AI to exploit their monstrous datasets to give clients the best insight. Three distinct gatherings use AI:

- Researchers
- Data researchers
- Programmers.

They would all be able to utilize the equivalent toolset to work together with one another and improve their effectiveness.

Google doesn't simply have any information; they have the world's most monstrous PC, so TensorFlow worked to scale. TensorFlow is a library created by the Google Brain Team to quicken AI and profound neural organization research.

It was designed to run on various CPUs or GPUs and even portable working frameworks, and it has a few coverings in a few dialects like Python, C++ or Java.

In this instructional exercise, you will learn

TensorFlow Architecture: Tensor stream engineering works in three sections:

- Pre-preparing the information
- Build the model
- Train and gauge the model

It is called Tensor stream since it accepts contribution as a multi-dimensional cluster, otherwise called tensors. You can build such a flowchart of activities (called a Graph) that you need to perform on that input. The info goes in toward one side, and afterward it moves through this arrangement of numerous tasks and comes out the opposite end as yield.

This is the reason it is called TensorFlow on the grounds that the tensor goes in its courses through top notch activities, and afterward it comes out the opposite side.

TensorFlow can equipment, and programming prerequisites can be grouped into

Improvement Phase: This is the point at which you train the mode. Preparing is typically done on your Desktop or PC.

Run Phase or Inference Phase: Once preparing is done Tensor stream can be run on a wide range of stages. You can run it on

- Desktop running Windows, macOS or Linux
- Cloud as a web administration
- Mobile gadgets like iOS and Android

You can prepare it on various machines then you can run it on an alternate machine, when you have the prepared model.

The model can be prepared and utilized on GPUs just as CPUs. GPUs were at first intended for computer games. In late 2010, Stanford analysts found that GPU was additionally awesome at network tasks and polynomial math so it makes them exceptionally quick for doing these sorts of

estimations. Profound learning depends on a great deal of framework increase. TensorFlow is quick at figuring the lattice increase since it is written in C++. Despite the fact that it is actualized in C++, TensorFlow can be gotten to and constrained by different dialects for the most part, Python.

At long last, a critical element of TensorFlow is the Tensor Board. The Tensor Board empowers to screen graphically and outwardly the thing TensorFlow is doing.

Rundown of Prominent Algorithms upheld by TensorFlow

- Linear relapse: `tf.estimator.LinearRegressor`
- Classification: `tf.estimator.LinearClassifier`
- Deep learning arrangement: `tf.estimator.DNNClassifier`
- Deep learning wipe and profound: `tf.estimator.DNNLinearCombinedClassifier`
- Booster tree relapse: `tf.estimator.BoostedTreesRegressor`
- Boosted tree characterization: `tf.estimator.BoostedTreesClassifier`

Anaconda Navigator: Anaconda Navigator is a constrictor Navigator is a work area graphical UI (GUI) remembered for Anaconda circulation that permits you to dispatch applications and effectively oversee conda bundles, conditions and channels without utilizing order line orders. Pilots can look for bundles on Anaconda Cloud or in a neighborhood Anaconda Repository. It is accessible for Windows, Macintosh OS and Linux.

Utilization of Navigator: In request to run, numerous logical bundles rely upon explicit adaptations of different bundles. Information researchers regularly utilize various forms of numerous bundles, and utilize different conditions to isolate these various renditions.

The order line program conda is both a bundle administrator and a climate supervisor, to help information researchers guarantee that every form of each bundle has all the conditions it requires and works accurately.

Pilot is a simple, point-and-snap approach to work with bundles and conditions without expecting to type conda orders in a terminal window. You can utilize it to discover the bundles you need, introduce them in a climate, run the bundles and update them, all inside Navigator.

Applications that can be access utilizing pilot: The accompanying applications are accessible naturally in Navigator:

- JupyterLab
- Jupyter Notebook
- QTConsole
- Spyder
- VSCode
- Glueviz
- Orange 3 App
- Rodeo
- RStudio

Progressed conda clients can likewise construct your own Navigator applications

Running code with Navigator: The most straightforward path is with Spyder. From the Navigator Home tab, click Spyder, and compose and

execute your code.

You can likewise utilize Jupyter Notebooks in a similar way. Jupyter Notebooks are an inexorably mainstream framework that consolidate your code, distinct content, yield, pictures and intelligent interfaces into a solitary journal document that is altered, seen and utilized in an internet browser.

New highlights in 1.9 are:

- Add uphold for Offline Mode for all climate related activities.
- Add uphold for custom design of fundamental windows joins.
- Numerous bug fixes and execution upgrades.

Information Cleaning

Information cleaning is a vital advance in information investigation. For instance, we generally check for missing qualities in the information by running `pd.isnull()` which checks for invalid Values, and returns a boolean cluster (a variety of valid for missing qualities and bogus for non-missing qualities). To get an amount of invalid/missing qualities, run `pd.isnull().sum()`. `pd.notnull()` is something contrary to `pd.isnull()`. After you get a rundown of missing qualities you can dispose of them, or drop them by utilizing `df.dropna()` to drop the lines or `df.dropna(axis=1)` to drop the sections. An alternate methodology is fill the missing quantities with different qualities by utilizing `df.fillna(x)` which fills the missing qualities with `x` (you can put there anything you desire) or `s.fillna(s.mean())` to supplant all invalid qualities with the (mean can be supplanted with practically any capacity from the insights segment).

It is once in a while important to supplant values with various qualities. For instance, `s.replace(1,'one')` would supplant all qualities equivalent to 1 with 'one'. It's conceivable to do it for numerous qualities: `s.replace([1,3],['one','three'])` would supplant each of the 1 with 'one' and 3 with 'three'. You can likewise rename explicit segments by running: `df.rename(columns={'old_name': 'new_name'})` or use `df.set_index('column_one')` to change the record of the information outline.

Join/Combine

The last arrangement of fundamental Pandas orders are for joining or consolidating information casings or lines/segments. The three orders are:

- `df1.append(df2)`— add the lines in `df1` to the furthest limit of `df2` (sections ought to be indistinguishable)
- `df.concat([df1, df2],axis=1)`—add the segments in `df1` to the furthest limit of `df2` (columns ought to be indistinguishable)
- `df1.join(df2, on=col1, how='inner')`—SQL-style join the sections in `df1` with the segments on `df2` where the columns for `col` have indistinguishable qualities. `how` could be equivalent to one of: 'left', 'right', 'external', 'inward'

NUMPY

Numpy is one such ground-breaking library for cluster preparation alongside a huge assortment of significant level numerical capacities to work on these exhibits. These capacities fall into classes like Linear Algebra, Trigonometry, Statistics, Matrix control, and so on

Getting NumPy: NumPy's principle object is a homogeneous multidimensional exhibit. Not at all like python's exhibit class which just handles one-dimensional cluster, NumPy's `ndarray` class can deal with multidimensional cluster and gives greater usefulness. NumPy's measurements are known as tomahawks. For instance, the exhibit beneath has 2 measurements or 2 tomahawks in particular lines and sections. In some cases measurement is otherwise called a position of that specific exhibit or lattice.

Bringing in NumPy: NumPy is imported utilizing the accompanying order. Note here `np` is the show followed for the nom de plume with the goal that we don't have to compose `numpy` without fail.

- `import numpy as np`

NumPy is the essential library for logical calculations in Python and this article delineates a portion of its most as often as possible utilized capacities. Understanding NumPy is the main significant advance in the excursion of AI and profound learning.

Sklearn

In python, scikit-learn library has a pre-fabricated usefulness under sklearn. Pre-preparing.

Next thing is to include extraction Feature extraction is a property decrease measure. Dissimilar to include choice, which positions the current ascribes as per their prescient importance, highlight extraction really changes the credits. The changed credits, or highlights, are straight mixes of the first ascribed. At long last, our models are prepared utilizing Classifier calculation. We use nltk . order module on Natural Language Toolkit library on Python. We utilize the named dataset accumulated. The remainder of our marked information will be utilized to assess the models. Some AI calculations were utilized to order pre-handled information. The picked classifiers were Decision tree, Support Vector Machines and Random woodland. These calculations are extremely mainstream in content grouping errands.

SEABORN

a) Data Visualization in Python

Information representation is the control of attempting to comprehend information by putting it in a visual setting, so that examples, patterns and relationships that may not in any case be recognized can be uncovered.

Python offers various incredible diagramming libraries that come loaded with heaps of various highlights. Regardless of in the event that you need to make intuitive, live or profoundly tweaked plots python has a superb library for you.

To get a little outline here are a couple of mainstream plotting libraries:

- Matplotlib: low level, gives loads of opportunity
- Pandas Visualization: simple to utilize interface, based on Matplotlib
- Seaborn: undeniable level interface, incredible default styles
- ggplot: in view of R's ggplot2, utilizes Grammar of Graphics
- Plotly: can make intelligent plots

In this article, we will figure out how to make essential plots utilizing Matplotlib, Pandas perception and Seaborn just as how to utilize some particular highlights of every library. This article will zero in on the sentence structure and not on deciphering the diagrams.

b) Matplotlib

Matplotlib is the most well known python plotting library. It is a low-level library with a Matlab like interface which offers heaps of opportunity at the expense of composing more code.

- To introduce Matplotlib, pip and conda can be utilized.
- pip introduce matplotlib
- conda introduce matplotlib

Matplotlib is explicitly useful for making essential diagrams like line graphs, bar outlines, histograms and some more. It very well may be imported by composing:

- import matplotlib.pyplot as plt

In Matplotlib we can make a line diagram by calling the plot strategy. We can likewise plot different sections in a single diagram, by circling through the segments we need, and plotting every segment on a similar hub.

4. CONCLUSION:

The horticulture area needs a colossal upgrade to endure the changing states of the Indian economy. Alongside the advances in machines and advances utilized in cultivating, helpful and precise data about various matters additionally assumes a critical part in it. Various occasional, monetary and organic examples impact the harvest creation however capricious changes in these examples lead to an extraordinary misfortune to ranchers. These dangers can

be diminished when reasonable methodologies are utilized on information identified with soil type, temperature, air pressing factor, mugginess and yield type. While, harvest and climate anticipating can be anticipated by getting valuable experiences from this agrarian information that guides ranchers to settle on the yield they might want to plant for the impending year prompting greatest benefit. This paper presents an overview on the different calculations utilized for climate, crop yield, and harvest cost expectation.

REFERENCES

- [1] G Rasul, Q. Z. Chaudhry, A. Mahmood, K.W. Hyder, "Effect of 28–40 Temperature Rise on Crop Growth & Productivity," *Pakistan Journal of Meteorology*, no. 15, vol. 8, pp. 7-8, 2011.
- [2] PrasadaRao.P.V.R.D, Vineetha.B, Priyanka.Ch and Satish.V, "A Comparative Study On Recommendation System Using Hybrid Approach," *International journal of Mechanical Engineering and Technology(IJMET)*, vol. 9, no. 1, January, 2018.
- [3] Balaraju.J and PrasadaRao.P.V.R.D, "Designing authentication for hadoop cluster using DNA algorithm," *International Journal of Recent Technology and Engineering*, 2019.
- [4] A. Mahato, "Climate Change and its Impact on Agriculture," *International Journal of Scientific and Research Publications*, Vols. 4, no. 4, no. April, ISSN 2250-3153, pp. 4-5, 2.14.
- [5] J. Kaur, "Impact of Climate Change on Agricultural Productivity and Food Security Resulting in Poverty in India," *Università Ca' Foscari Venezia*, vol. 23, pp. 16-18, 2017.
- [6] Pratap S. Birthal, Md. Tajuddin Khan, Digvijay S. Negi, Shaily Agarwal, "Impact of Climate Change on Yields of Major Food Crops in India: Implications for Food Security," *Shaily Agarwal Agricultural Economics Research Review*, vol. 27, pp. 145-155, 2014.
- [7] S. Reinhard. J.P. Powell, "Measuring the effects of extreme weather events on yields in Weather and Climate Extremes," *Elsevier*, vol. 12, pp. 69-79, 2016.
- [8] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159-175, 2013.
- [9] B. Dumont, V. Leemans, Salvador Ferrandis, Bernard Bodson, Jean-Perrie Destain, "Assessing the potential of an algorithm based on mean climatic data to predict wheat yield," *Precision Agriculture*, vol. 15, no. June, pp. 255-272, 2014.
- [10] B Basso, B Bodson, V. Leemans, B. Bodson, J-P Destain, M-F Destain, "A comparison of within season yield predictions algorithm based on crop model behaviour analysis," *Agricultural and Forest Meteorology*, vol. 204, no. May, pp. 10-21, 2015.
- [11] Betty. J, Shem G Juma, Everline. O, "On the Use of Regression Models to Predict Tea Crop Yield Responses to Climate Change: A Case of Nandi East Sub-County of Nandi County Kenya," *Assessing the Value of Systematic Cycling in a Polluted Urban Environment Climate*, Vols. 5, no. 3, no. July, p. 5, 2017.
- [12] Christian Baron, Mathieu Vrac, Oettli. P, Sultan. B, "Are regional climate models relevant for crop yield prediction in West Africa," *Environmental Research Letters*, vol. 6, pp. 2-6, 2011.