IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

ENHANCED FARMING USING ML BASED ON MACHINE LEARNING

DR.Gurumoorthy (Assistant Professor) Department of Computer Science and Engineering
St. Peters Engineering College, Dullapally, Maisammaguda, Medchal,Hyderabad,Telangana 500043

A. Sahith Goud
Department of Computer Science and Engineering
St. Peters Engineering College, Dullapally, Maisammaguda, Medchal,Hyderabad,Telangana 500043

A
Muda Yamini
Department of Computer Science and Engineering
St. Peters Engineering College, Dullapally, Maisammaguda, Medchal,Hyderabad,Telangana 500043

P. Sai Pravalika
Department of Computer Science and Engineering
St. Peters Engineering College, Dullapally, Maisammaguda, Medchal,Hyderabad,Telangana 500043

ABSTRACT:

Earlier, crop cultivation was undertaken on the basis of farmers' hands-on expertise. However, climate change has begun to affect crop yields badly. Consequently, farmers are unable to choose the right crop/s based on soil and environmental factors, and the process of manually predicting the choice of the right crop/s of land has, more often than not, resulted in failure. Accurate crop prediction results in increased crop production. This is where machine learning plays a crucial role in the area of crop prediction. Crop prediction depends on the soil, geographic and climatic attributes. Selecting appropriate attributes for the right crop/s is an intrinsic part .Machine learning provides accurate prediction and estimation of farming parameters to optimize the economic efficiency .Enhanced farming aims to make predictions by better use of agricultural resources such as rainfall, humidity and other biological resources like ph value ,potassium etc.

KEYWORDS: Machine Learning, Crop Prediction ,Classification ,Soil management

INTRODUCTION:

The influence of climate change and its unpredictability, has caused the majority of the agricultural crops to be affected in terms of their production and maintenance. Forecasting or predicting the crop yield well ahead of its harvest time would assist the strategists and farmers in taking suitable measures for selling and storage. Accurate prediction of crop development stages plays an important role in crop production management. Agriculture plays a critical role in the global economy. Pressure on the agricultural system will increase with the continuing expansion of the human population. Agri-technology and precision farming, now also termed digital agriculture, have arisen as new scientific fields that use data intensive approaches to drive agricultural productivity while minimizing its environmental impact. The data generated in modern agricultural operations is provided by a variety of different sensors that enable a better understanding of the operational environment (an interaction of dynamic crop, soil, and weather conditions) and the operation itself (machinery data), leading to more accurate and faster decision making.

Machine learning (ML) has emerged together with big data technologies and high-performance computing to create new opportunities to unravel, quantify, and understand data intensive processes in agricultural operational environments. Among other definitions, ML is defined as the scientific field that gives machines the ability to learn without being strictly programmed Year by year, ML applies in more and more scientific fields including, for example, bio informatics , bio chemistry etc. Predictions will also support the allied industries for strategizing the logistics of their business. Several means and approaches of predicting and demonstrating crop yields have been developed earlier with changing rates of success, as these don't take into consideration the weather and its characteristics and are mostly empirical. The yield prediction is still considered to be a major issue that remains to be explained based on available data for some agricultural areas.

LITERATURE REVIEW:

Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving. The ideal characteristic of artificial intelligence is its ability to rationalize and take actions that have the best chance of achieving a specific goal. A subset of artificial intelligence is machine learning, which refers to the concept that computer programs can automatically learn from and adapt to new data without being assisted by humans. Deep learning techniques enable this automatic learning through the absorption of huge amounts of unstructured data such as text, images, or video. Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves.

Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks. A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers; but not all machine learning is statistical learning. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. In its application across business problems, machine learning is also referred to as predictive analytics. The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers to learn automatically without human intervention or assistance and adjust actions accordingly.

EXISTING SYSTEM:

Use of weather forecasting:

With the change in climatic condition and increasing pollution it's difficult for farmers to determine the right time for sowing seed, with help of Artificial Intelligence farmers can analyse weather conditions by using weather forecasting which helps they plan the type of crop can be grown and when should seeds be sown.

Soil and crop health monitoring system:

The type of soil and nutrition of soil plays an important factor in the type of crop is grown and the quality of the crop. Due to increasing, deforestation soil quality is degrading and it's hard to determine the quality of the soil.

Analyzing crop health by drones:

Sky Squirrel Technologies has brought drone-based Ariel imaging solutions for monitoring crop health. In this technique, the drone captures data from fields and then data is transferred via a USB drive from the drone to a computer and analyzed by experts.

Precision Farming and Predictive Analytics:

AI applications in agriculture have developed applications and tools which help farmers inaccurate and controlled farming by providing them proper guidance to farmers about water management, crop rotation, timely harvesting, type of crop to be grown, optimum planting, pest attacks, nutrition management.

PROPOSED SYSTEM:

Predicting suitable crops for cultivation is an essential part of agriculture, with machine learning algorithms playing a major role. The vast datasets obtained can be used for crop prediction on a massive scale. In this project we use Decision Tree, Random Forest Tree Algorithms for predicting crop suitable for increasing crop production depending on the soil, geographic and climatic attributes.

Enhanced farming aims to make predictions by better use of agricultural resources such as rainfall, humidity and other biological resources like ph value ,potassium etc.

DATASET DESCRIPTION:

The dataset parameters used to predict the crop name is as follows:

1.Ph value - Most soils have pH values between 3.5 and 10. In higher rainfall areas the natural pH of soils typically ranges from 5 to 7, while in drier areas the range is 6.5 to 9. Soils can be classified according to their pH value: 6.5 to 7.5—neutral.

2.Rainfall - Soil is also greatly affected by rainfall. If it is too wet or too dry, nutrients in the soil can run off and not make it to the plants' roots, leading to poor growth and overall health.

3.Humidity - Humidity is important to make photosynthesis possible. Good humidity around the plant is even more important than for most crops.

4.Temperature - Soil temperature directly affects plant growth. In other words, nearly every crop slows down its growth when soil temperatures are below 90 C and above 50C.

5.Potassium - Potassium chemical symbol (K) is one of 17 essential nutrients required for plant growth and reproduction. potassium improves the overall health of growing plants and helps them fight against disease, it is known as the "quality" nutrient.

7.Nitrogen - Nitrogen is the main limiting nutrient after carbon, hydrogen and oxygen for photosynthetic process, phyto-hormonal, proteomic changes and growth-development of plants to complete its life cycle.

8.Phosphorus - Phosphorus (P) is an essential nutrient required for crop and animal production and for human health. ... For over 50 years, farmers in developed countries have been encouraged to invest in P fertilizers and to improve soil P fertility to maximize crop output

IMPLEMENTATION:

Sample crop data used is as shown below

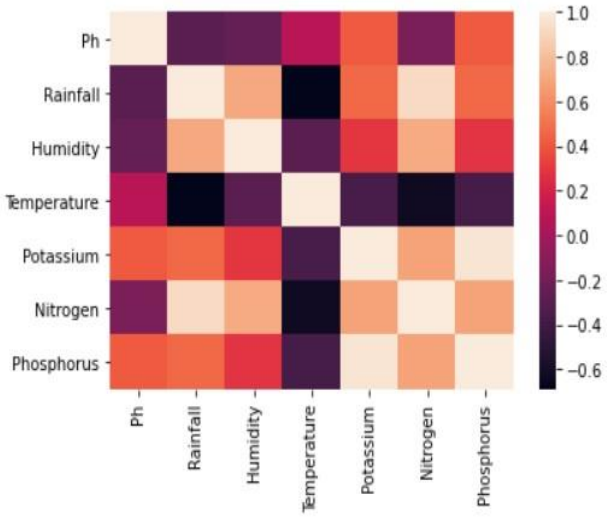
Ph	Rainfall	Humidity	Temperature	Potassium	Nitrogen	Phosphorus	Crop
4	146	50	23	36	95	34	paddy
4	168	52	24	31	107	30	paddy
6	57	19	29	37	51	29	mustard plant
4	169	59	25	34	112	29	paddy
6	185	67	31	40	121	39	paddy
7	155	47	24	80	150	77	sugarcane
5	31	33	41	21	47	19	wheat
4	33	40	32	18	46	21	maize
5	39	40	41	23	50	23	wheat
4	41	23	32	27	46	28	mustard plant
4	30	37	41	21	50	22	wheat
5	38	40	43	22	50	22	wheat
7	160	46	22	75	142	79	sugarcane
7	58	48	22	30	56	30	maize
4	43	26	34	25	45	22	mustard plant
4	37	36	34	22	45	18	maize
6	48	50	33	25	53	24	wheat

EXPLORATORY DATA ANALYSIS:

Exploratory Data Analysis(EDA) refers to the critical process of performing initial investigations on data so as to discover patterns, to spot anomalies, to test hypothesis and to check assumptions with the help of summary statistics and graphical representations. In statistics, exploratory data analysis is an approach of analyzing data sets to summarize their main characteristics, often using statistical graphics and other data visualization methods. EDA also helps stakeholders by confirming they are asking the right questions. EDA can help answer questions about standard deviations, categorical variables, and confidence intervals.

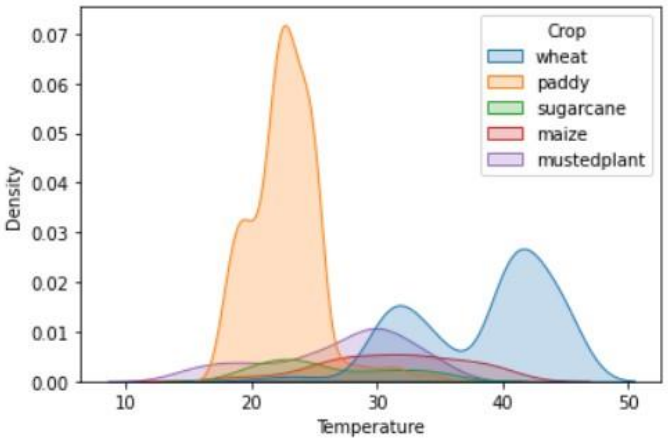
CORRELATION MATRIX:

A correlation matrix is simply a table which displays the correlation. The measure is best used in variables that demonstrate a linear relationship between each other.

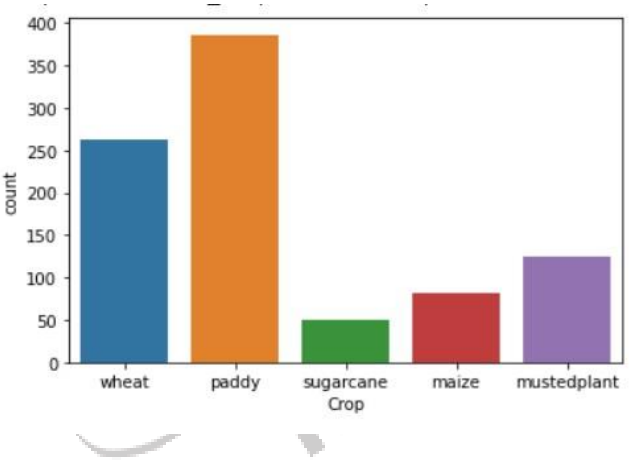


KDE PLOT:

Temperature vs Crop

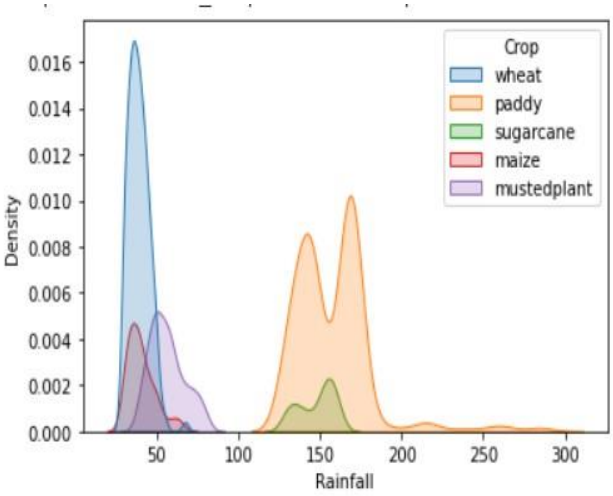


COUNT PLOT:



KDE PLOT:

Rainfall vs Crop



PREDICTIVE ANALYSIS:

Machine learning is a form of predictive analytics that advances organizations up the business intelligence (BI) maturity curve, moving from exclusive reliance on descriptive analytics focused on the past to include forward-looking, autonomous decision support. Predictive analytics is the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data. Predictive analytics encompasses a variety of statistical techniques (including machine learning, predictive modelling and data mining) and uses statistics (both historical and current) to estimate, or 'predict', future outcomes. Predictive analytics involves certain manipulations on data from existing data sets with the goal of identifying some new trends and patterns.

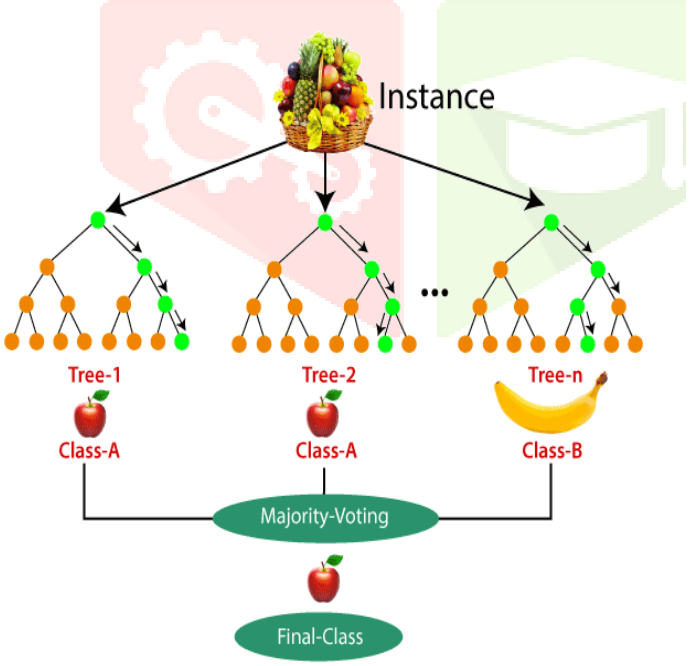
RANDOM FOREST TREE :

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML.

Random Forest is a classifier that contains a number of **decision trees** on various subsets of the given dataset and takes the average to improve the **predictive accuracy** of that dataset.

Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

The **greater number of trees** in the forest leads to **higher accuracy** and prevents the problem of **overfitting**.



A RANDOM FOREST TREE CLASSIFIER:

Predictive analysis

```
[19] X=df[['Ph', 'Rainfall', 'Humidity', 'Temperature', 'Potassium', 'Nitrogen',
           'Phosphorus']]
     Y=df['Crop']
```

Splitting Data

```
[20] from sklearn.model_selection import train_test_split
```

```
[21] X_train, X_test, y_train, y_test = train_test_split(X,Y,test_size=0.2)
```

Random Forest

```
[23] from sklearn.ensemble import RandomForestClassifier
     model = RandomForestClassifier(n_estimators=20)
     model.fit(X_train, y_train)
```

```
RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight=None,
                       criterion='gini', max_depth=None, max_features='auto',
                       max_leaf_nodes=None, max_samples=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=20,
                       n_jobs=None, oob_score=False, random_state=None,
                       verbose=0, warm_start=False)
```

DECISION TREE:

Decision Tree

```
[28] from sklearn import tree
     dt=tree.DecisionTreeClassifier()
     X=df[['Ph', 'Rainfall', 'Humidity', 'Temperature', 'Potassium', 'Nitrogen',
           'Phosphorus']]
     Y=df[['Crop']]
```

```
dt.fit(X_train,y_train)
```

```
DecisionTreeClassifier(ccp_alpha=0.0, class_weight=None, criterion='gini',
                       max_depth=None, max_features=None, max_leaf_nodes=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, presort='deprecated',
                       random_state=None, splitter='best')
```

```
[31] dt.score(X_test,y_test)
```

```
0.9834254143646409
```

```
[32] dt.predict([[4,30,37,21,23,49,20]])[0]
```

```
'maize'
```

CONCLUSION:

For generations, the production of essential food crops has been correlated with agriculture. In agriculture, crop cultivation prediction is a key factor. Generally, agro-climatic input parameters such as soil texture, rainfall, and temperature influence crop production. Input parameters for agriculture vary from region to region, and it is daunting to collect such information over large tracts of land. The vast datasets obtained can be used for crop prediction on a massive scale. Predicting suitable crops for cultivation is an essential part of agriculture, with machine learning algorithms playing a major role in such prediction in recent years. In this era of technology and data science, the agricultural sector stands to benefit greatly from properly implemented techniques. There are three common machine learning techniques: supervised, unsupervised, and reinforcement learning. This work uses supervised learning

REFERENCES:

[1]https://www.innovationnewsnetwork.com/machine-learning-in-agriculture/9892/

[2]https://www.analyticsvidhya.com/blog/2020/11/artificial-intelligence-in-agriculture-using-modern-day-ai-to-solve-traditional-farming-problems/

[3]https://www.cropscience.bayer.com/innovations/data-science/a/machine-learning-uses-agriculture

[4]https://www.kdnuggets.com/2019/05/machine-learning-agriculture-applications-te chniques.html

[5]https://www.sciencedirect.com/science/article/pii/S25897217203001 2X

[6]https://towardsdatascience.com/predicting-crops-yield-machine-learning-nanodegree-capstone-project-e6ec9349f69