We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS

BOOK CITATION INDEX

INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Precision in Agriculture Decision Making Based on Machine Learning

Ratnmala Bhimanpallewar and M.R. Narasinga Rao

Abstract

Farming is the one of the major occupations in India. Increase in population is increasing the demand of food, whereas soil degradation causing decrease in yield. Technology is contributing in agriculture domain through software and hardware enhancement. One of the software-based contribution is for predicting the suitable crop. Same field can be suitable for one crop and not for another one, so it is better to choose the one which can lead to better yield. There are many predictive algorithms available. Algorithms which can work for suitability analysis need to test and choose the best one. Such predictive algorithms need dataset in appropriate format. Once the quality data is available correct predictions can be made. Data mining, machine learning are the branches comprise of algorithms, which can be trained based on dataset. Here we are introducing algorithms for decision making based on field data.

Keywords: machine-learning, agriculture, suitability, yield, dataset, algorithm, fragmented land, crop

1. Introduction

Agriculture is the main occupation in India and no doubt experienced farmers are the expert at doing everything. They can take the decisions; some decisions may fail so need to do some analysis for good results. One of the decisions we are considering here is "which crop to be adopted to have better yield?". It needs to analyze all the features on which yield is dependent and predict the suitable crop/crops. The basic feature i.e. agriculture land is the natural feature, which cannot prepare artificially and that too increase in the soil degradation is one of the serious issues farmers are facing [1]. There are multiple reasons affecting on soil quality one of them is, use of fertilizers without knowledge. For increasing yield farmers are using unbalanced quantity of fertilizers [2], without knowing the effect of it.

Even young farmers cannot take decision about which crop is suitable for current situation (soil quality, environment etc.). Farmers are adopting some crop and using chemicals for better yield, we have a good example for this from case study at Kolhapur district [3] which was resulted in causing cancer. To avoid such worst impacts we are thinking about recommending the suitable crop/crops. To get the genuine decision we are using historical data. Once the pattern of historical results identified successfully, predictions can be done appropriately.

There are many algorithmic approaches available for predictive analysis, machine learning is the branch which makes computer to learn from the dataset available. We are discussing about the 1) Role of machine learning in precise decision making about agriculture, 2) How to classify suitability level of crop/crops and feature values.

2. Agriculture decision making: Current scenario

Since old days farmers are taking farming decisions about agriculture by experience. Here we are discussing about the decision of choosing the crop to be adopted. Though agriculture is a major sector in India, same trend is going on. Automated and advanced systems in the form of software and physical machines are invented and available for use in developed countries, some of them are available in developing countries like India as well. Available facilities are either unaffordable or not easily approachable. Some systems are developed for some particular geographical area, it cannot be adopted for India as it is. Software techniques available in developed countries are not applicable due to fragmented land, because those systems are developed considering unfragmented land.

In India Krishi Vigyan Kendra (KVK) are the centers made available for farmer's guidance. It is responsible to spread technical knowledge among farmers. They conduct training, awareness programs to achieve some goals set as below:

1. Develop advisory services for farmers.

2. Conduct training program on different trends.

3. Conduct training programs for different level of people working in agriculture.

4. Agriculture field testing

5. Demonstration for recent innovations

It is playing an important role for needy farmers. Multiple KVK centers are there, so they can work according to the location to develop location specific solutions. It also has contribution in providing quality products like planting material, seed, organic material, livestock related products etc. As above mentioned, all agriculture related centers are available to facilitate people working in agriculture sector.

Despite all the knowledge centers, testing facilities available, farmers are continuing with traditional methods. The major change they have adopted is after Green revolution. Green revolution has shown drastic increase on yield [4, 5]. Other than using high-yielding varieties of seeds and the improved quality of fertilizers farmers are not interested to adopt new technics. There are many reasons behind it, as discussed further.

The one who work in farm are not well educated, not aware about the ongoing innovations. Even if they become aware, they cannot afford it. Economic condition of most of the farmers in India is not sufficient to purchase the automated systems available in the market. At another end, the educated, economically well, and aware people not much involved in actual farming and related occupations. The traditional farming methods are dependent on natural parameters. Uncertainty in the nature directly effects on agriculture production, so not getting yield as expected.

Thus, since long ago there is no improvement in the economic condition of small/marginal scale farmers in India.

3. Existing systems

Decision making system available in other countries. One of the popular decision making system is Agriculture Land Suitability Evaluator (ALSE), is the crop specific evaluator at Peninsular Malaysia [6]. Here crop specific means, it works for mango, citrus, guava, papaya and banana as well. Base for decision making is cultivation history, cultivation knowledge, land characteristics, climate features such as annual precipitation, dry season length per month, land slope, nutrient availability, and retention. Some other features are used for land availability and suitability evaluation in Tuban Regency, Java Island [7]. Spatial multi-criteria analysis has been done based on parameters like, land elevation, slope, slope direction, land use/land cover, land capability, integrating soil order, climate, and accessibility. Outcome prepared was land use plan by the spatial pattern. This is the area having most fertile land in the country. The weights are assigned to above mentioned criteria with the help of eight experts involved for sub-criterions. Criteria wise scores are assigned in the range of 0–10, according to involvement of sub criterions under each criterion. Weighted sum overlay method used those weights to prepare suitability map. Suitability analysis for crop Soybean in Indonesia [8], to satisfy the local need. Regular domestic consumption of soybean was more so the, need were not fulfilled. The research was conducted in Karawang Regency, West Java, Indonesia to identify suitable area for soyabean plantations in paddy fields and prepared plan for it. The suitability classes defined according to FAO categories from suitable (S2) to not suitable (N). For wheat crop suitability analysis conducted in North Carolina [9]. The case study under taken was rain-fed wheat. Five criterions considered were soil-fertility, climate, soil-features, soil-organic-matter, soil-quality, soil-chemistry and seventeen sub-criteria under that. This system also considers geographic information systems (GIS) as base and the square root method is used, called multi-criteria analysis. Percentage of land suitability for organic wheat was highly suitable- 18.6% and moderately suitable- 76.8% in Duplin country. Existing yield simulation method was also based on Moderate Resolution Imaging Spectroradiometer [10]. A case study for corn and soybean yield simulated within a certain scope of area, predictions are given by the United States Department of Agriculture (USAD)- National Agricultural Statistics Service (NASS). All above discussed systems are suitable for unfragmented land not for fragmented lands I India.

Decision making system available in India

As discussed above still the decision making in India mainly for small/marginal scale fields is by traditional way. So, farmers are not getting expected yield due to many reasons like decreasing quality of soil, uncertainty in nature. Till the moment automated machines are not used by small/marginal scale Indian farmers and are dependent on labors. There are many reasons which makes them to be dependent on labors as below:

1. Poor economic condition, so unable to purchase machines.

2. In fragmented lands, it is difficult to use the big machineries like tractor and tractor accessories.

3. Unaffordable cost of automated systems.

4. Unaware due to lack of interest.

Though the farming is the main occupation in India, due to urbanization people are moving to urban areas in search of jobs and it causes labor deficiency. So, labor cost is increasing which add on to farming expenses and again it is lowers the economic condition of farmers.

4. Introduction to technology in agriculture decision making

Decision making is in the context of choosing the appropriate crop which give good outcome in available natural conditions like soil, environment. The reason behind it is soil and environment cannot be controlled. So, first we need to study about the technologies available to analyze the data collected through monitoring and recording the soil and environmental features. There are existing prediction techniques we can study to choose the suitable method and we can do the crop suitability analysis based on historical data. Suitability analysis will give the level of suitability for crop/crops and then user can take the decision accordingly. The crop/crops having fair suitability level can be adopted for better yield.

We have divided this advisory system into sub parts as below:

4.1 Collecting data

4.1.1 Identifying the features

The first important task is to identify the features affecting on crop yield. As per the guidelines available at country level [11], state-wise [12], local [13] through variety of sources. We have listed the fallowing recurrent features from variety of sources and suitable ranges of the possible features.

a. Topography

b. Temperature

c. Soil type

d. Soil Quality

e. Rainfall

f. Soil moisture

g. pH

h. EC

i. Soil nutrients: Nitrogen, Phosphorus, Potassium

j. Soil micronutrients

All above features and some features can be added or removed as per requirement and availability of data like humidity, soil texture etc. Some of the features remain constant for the period of a year or more than that ex. Topography of land does not change for years. Some features need to measure for the period of a season or less than that ex. rainfall required at the beginning that is seeding phase of the crop is different than the growing phase of the crop. As usual we are considering two seasons of cropping i.e. Kharif and Rabi. There are three main phases of the crop seeding, growing and maturity. Each phase has different requirement for rainfall, soil-moisture, temperature, nitrogen level, phosphorus level and potassium level. We will discuss about the influence of features on crop growth one by one. To consider these values for analysis purpose we need to normalize the feature values. Some of the feature values are considered in the numeric form and some need to convert into categorical form. Textual values cannot be considered as it is, it needs to map to the numeric categories. For better understanding of the technique let us take an example of the crop wheat.

a. Topography

Topography is nothing but the slope of the crop land. The field having slope less than 10 degree is good to use for cropping. It is a natural feature which can be controlled artificially. With or without analysis we can say this is vital feature contributing towards decision making for crop suitability. Sloppy fields are not good at holding the water at higher ends and may have clayey soil at lower end this uneven nature is not suitable for crop growth. Some of the crops with less water requirement can be adopted provided the slope is less. To use the topography values for algorithm purpose we have categorized it. Plane surface is always most suitable for every crop so, the categorization is done as 1-Plane, 2- slope less than 10 degree and 3-slope more than 10 degree.

b. Temperature

It is again a natural parameter. Geographical location of India is such that enough solar energy is available. This feature values need to consider phase wise. Suitable ranges of temperature is discussed in **Table 1**. If temperature goes below or above the range crop yield get reduced. The values of temperature can be considered as it is in the numeric form with unit °C (degree Celsius).

c. Soil type

Type of the soil plays a key role in crop yield. Water holding capacity is dependent on the soil type. For some of the crops clayey soil is good and for

Sr. No.	Temperature	Soil quality	Suitability
1	24°C	Good	Level1
2	22°C	Good	Level1
3	20 °C	Good	Level1

Table 1.
Sample tuple in dataset.

some loamy. Few crops can be grown in sandy soil as well. Soil type is textual value, cannot be considered as it is. Here we are categorizing it according to water holding capacity. First category is 1- Loamy, 2- Clayey, 3- Slity, 4- Sandy.

d.Soil quality

Quality of the soil is dependent on multiple features like soil nutrients, micronutrients, texture, water holding capacity and it may varies according to the chemical used, erosion occurred etc. Krishi Vigyan Kendra helps to know the quality of soil in terms of levels. With reference to that we are considering the soil quality levels as 1- Good, 2- Moderate, 3- Marginal 4- Low.

e. Rainfall

Rainfall is the natural feature on which other features dependent. Water requirement for different crop is different. Precipitation has some annual pattern, sometimes it may vary. Here we can directly consider the numeric values or range of it for analysis purpose. This feature needs to measure according to phases of the crop.

f. Soil moisture

Water holding capacity of the soil, supplied water decides the soil moisture level. Good quality and type of soils has enough moisture content. Not only rainfall but irrigation sources also contribute to decide the soil moisture values. Here we are taking its numerical values as it is phase wise.

g.pH

It shows acidity or alkalinity of soil and is measured in pH units. Different crops can bear different level of acidity in soil and water. This numerical value is considered in the range of 0 to 14.

h.EC

Soil electrical conductivity is measure of amount of salinity, one of the indicators of soil health. Excess salinity levels occur in arid and semiarid regions. For this feature as well numeric values considered directly. The range of it is from 0.611 to 25.9 dS m − 1.

i. Soil nutrients: Nitrogen, Phosphorus, Potassium

Major soil nutrients are nitrogen, phosphorus, potassium. This feature is artificially manageable. Deficiency can be resolved by adding the fertilizers available in the market but excess amount available can not reduced in any way. If farmer get their soil tested, they come to know the existing level of nutrients and amount of fertilizers to be added. But they do not approach for it and add the fertilizers without knowing the requirement. This is leading to soil degradation. Once we know the suitable crop and existing amount of nutrients, it is easy to recommend the appropriate amount of fertilizers to add and avoid the soil degradation up to certain extent.

j. Soil micronutrients

There are subcategories of soil micronutrients. Soil cab be tested into laboratory to know the availability of micronutrients, so that we can understand one aspect of soil quality and take future decision.

These are some basic feature we have considered here. We need to keep all the information, while recording these feature values: 1) what was the location? 2) what were the date and time? 3) what was the cropping season? 4) what was the crop/crops adopted? and 5) what was the final yield? There are some other features affecting on yield but either has less effect or values are not available easily. If more factors are considered, results get quite improved.

4.1.2 Collecting data using feature values

There are two different sources from which dataset can be collected. Either we can have sensor-based device which will sense actual field features and record the dataset. According to features we want to monitor; device can be integrated with respective sensors (available in market) and microcontroller to control the data recording [14] and storing process. This method can be referred as monitoring feature values. Another method is to prepare dataset from available online or offline sources called as gathering feature values.

4.1.2.1 Monitoring feature values

Environmental data is available with the meteorological department since long ago. Parameters required for agriculture are with the different scope ex. Rainfall measured by meteorological department is area wise, city wise whereas for agriculture purpose rainfall need to measure at specific location. Soil features also vary one farm to another farm as per the crop adopted and fertilizers used. So the global reports prepared cannot be referred as it is for deciding suitable crop. Crop specific monitoring system can be used as described below to measure the field specific features.

A monitoring system with microcontroller and accessories [14] to sense the nearby features like.

a. Rainfall sensor

b. Moisture sensor

c. pH sensor

d. EC sensor

e. Environmental sensor

f. NPK measuring kit

All the accessories like above can be used to monitor the actual field. There is lot of variety in the accessories available in the market. So according to the device used, scope of monitoring that feature get varies. Depending on the area of field one or more than one system needs to plant in the same field, so that whole farm will get

monitored to get more accuracy in the feature value. According to components used cost of this monitoring system get varies. The frequency of feature monitoring can be every second, minute or hour as per the code written, thus it is always editable as per user need. Data can be gathered in the tabular form, so that it can easily converted to SQL database for further analysis. Backup of data can be fetched from device as an when required. As we know microcontroller like raspberry-pi has own memory in the form SD card from which data can be fetched or we can make the provision for data transfer by using some network protocol, as raspberry-pi has default support for Wi-Fi, we just need to do the configuration and coding accordingly. This kind of data backup need only once in a season. It is better to maintain the data along with date and time (time is optional in some cases) and exact place of the data recording.

4.1.2.2 Gathering feature values

Metrological department already has the system to monitor the environmental features accurately. Also, Krishi Vigyan Kendra available at different places providing the values of soil and water features after soil and water testing. The drawback of this method is we do not get the data belongs to same location for which we want to do the suitability analysis. Metrological department has their setup at certain places only not at every place we might use for suitability analysis. KVK reports are also from the different fields, we cannot guarantee that all the soil samples collected are belongs to same location for which we want to do the suitability analysis. Here definitely we can filter out and choose the dataset belongs to same region which will give appropriate prediction. So, we collect the available and authorized data from metrology department, Krishi Vigyan Kendra and the offices working under agriculture universities and government maintaining data related to farming. This data is compiled further to fetch the values of identified features under Section 4.1.1 and crop yield value. Data need to maintain in tabular format so it can easily map to SQL database for further analysis. It is mandatory to maintain the data along with date and time (time is optional in some cases) and exact place of the data recording.

4.1.3 Preparing dataset from feature data

The frequency of monitored features can be changed by averaging the value for required interval. Same data can be customized for different purposes like weather prediction, crop suitability analysis, to understand the pattern of field features etc. Here we are talking about crop suitability analysis. Cropping seasons play important role in the crop specific suitability analysis. So, the interval for data analysis is also a season. Under season every crop goes through the three main phases seeding(s), growing(g), maturity(m). Each phase needs some basic requirement like in the seeding phase favorable temperature for wheat is 20°C -25°C, in the growing phase it is 15–30°C and in maturity phase it is 14°C -15°C [12] at state Punjab in India whereas the it can tolerate the temperature in the range of 3.5–35°C. It means if the crop gets that feature values in expected range in all the phases it will lead to better yield (i.e. high suitability level S1 or S2) otherwise yield may get reduced than expected (i.e. low suitability level S3 to N2). As per the crop season and the phase, interval of data collected is decided. Some of the features need to consider phase wise (ex. Rainfall) and for some features single value need to consider (ex. nitrogen, phosphorus, potassium (NPK). Along with the above identified features under Section 4.1.1 we also need to get to know the season, yield and crop/crops adopted for the fields considered through data gathering.

4.2 Analyze data for decision making

Collected data is further analyzed to understand the crop specific suitability for variety of crop for particular season, the one which is more suitable is advised to adopt. Here the logic is- based on historical data collected. From environmental data we understand the pattern of environmental features for current season and current phase of crop and predict that aggregate/common values in the respective season. Some of the features like NPK, pH are available as the current value of that feature, so it can be taken as it is not any prediction required here. Standard feature values/ranges required for crop is already well known to the farmers who are in farming business itself. For new farmers, the guidelines are available for cropping, feature values suitable are mentioned crop wise and phase wise under the guidelines available [12, 13]. Guidelines are available in the form of books, reports, online resources, also offline centers like KVK are always available to guide. The information provided by such resources is static information.

Here we are discussing about analyzing the suitability level of crop based on current situation dynamically. For such decision making, historical data with few current feature values can make available as discussed under Section 4.1. The suitable methodologies need to choose and customize for agriculture application. Suitability results need to compare with the existing results available, that is called as testing in machine learning. Variety of methods can be applied and then the methodology which gives more accuracy can be used further. Once the data is available, the methods based on data analysis are more suitable. Extracting the required information from available dataset is called as Data mining.

Data analysis means, identifying something based on current and historical data. Broadly it is categorized into two categories qualitative and quantitative analysis. If you want to get the answer for why, what, or how we need to go for qualitative analysis. If you want to know some statistical or categorical value, then go for quantitative analysis. Here we are discussing about suitability level analysis. According to Food and Agriculture Organization of the United Nations, suitability level is a categorical value [15], so quantitative approach has been used. Quantitative approaches further categorized as text analysis, statistical analysis, diagnostic analysis, predictive analysis, and prescriptive analysis. Here we want to predict the suitability based on historical and some current values, so need predictive analysis. If we have the historical database of the farm features along with the yield, yield of the crop/crops can decide respective suitability level of the crop for storing into dataset.

Ex. For crop C in particular farm area for season S, expected highest yield is Y_H Tones/Hector and lowest expected yield is Y_L Tones/Hector (it can be zero Tone in worst condition).

Then, interval between two adjacent levels Y_i calculated as,

$$Y_i = Y_H - Y_L / 4 \tag{1}$$

Five ranges of yield from higher to lower, are computed as,

$$\text{Level1 is} >= Y_H \tag{2}$$

$$\text{Level2 is} \langle Y_H \text{ and} \rangle = Y_H - Y_i \tag{3}$$

9

$$\text{Level3 is} \langle \text{Level2 and} \rangle = \text{Level2} - Y_i \tag{4}$$

$$\text{Level4 is} \langle \text{Level3 and} \rangle = \text{Level3} - Y_i \tag{5}$$

$$\text{Level5 is} < \text{Level4} \tag{6}$$

Example,
Expected high yield Y_H = 20 Tones/Hector and lowest yield Y_L = 3 Tones/Hector.
So,
Interval.

$$Y_i = Y_H - Y_L / 4 = 20 - 3 = 17 / 4 = 4.25$$

$$\text{Level1 is} >= 20 \, \text{Tones / Hector}$$

$$\text{Level2 is} \langle 20 \, \text{and} \rangle = Y_H - Y_i = 20 - 4.25 \, (15.75 \, \text{Tones / Hector})$$

$$\text{Level3 is} \langle 15.75 \, \text{and} \rangle = 15.75 - 4.25 \, (11.5 \, \text{Tones / Hector})$$

$$\text{Level4 is} \langle 11.5 \, \text{and} \rangle = 11.5 - 4.25 \, (7.25 \, \text{Tones / Hector})$$

$$\text{Level5 is} < 7.25 \, \text{Tones / Hector}$$

There are five suitability levels defined by FAO [15] as below,
S1- suitable.
S2- moderately suitable.
S3- marginally suitable.
N1- not suitable due to physical reasons.
N2- not suitable due to economic reasons.
Now if we know actual yield Y then, suitability level decided as,

$$\text{If Yield falls under Level1 then, suitability level} -> S1 \tag{7}$$

$$\text{If Yield falls under Level2 then, suitability level} -> S2 \tag{8}$$

$$\text{If Yield falls under Level3 then, suitability level} -> S3 \tag{9}$$

$$\text{If Yield falls under Level4 then, suitability level} -> N1 \tag{10}$$

$$\text{If Yield falls under Level5 then, suitability level} -> N2 \tag{11}$$

Thus, we can get the values for suitability level for historical dataset.

5. Machine learning: Methods and applications

Machine learning is the analysis technique which enables computer to learn from experience. Here experience is in the form of dataset. More the accuracy in dataset more accurate the results are. Here accuracy has the different dimensions. It depends on multiple things like 1)Duration of the dataset captured (ex. Data gathered from last 10 years is better than the data captured for last 1–2 years), 2) Features considered (It's better to consider maximum characteristics on which crop yield is dependent), 3)Frequency of the data recording (ex. The values of the features in field varies after certain time, so it's better if we could record every variation in dataset), 4)Duration of the data considered for analysis (ex. To do predictive analysis for the crop in kharif season we need to consider data recorded in kharif season only, the data recorded in rabi season will work as noise). 5)Missing data (if missing data is more in proportion then accuracy in data decreased). We can prepare the dataset from historical data available and the monitoring device. Learning from the historical data and comprehending for current situation is known as supervised machine learning, one of the best predictive analysis techniques. Supervised machine learning is further categorized into regression, classification, naive Bayesian model, random forest model, neural networks, support vector machines.

Classification is the method where we divide dataset into defined classes. Class is nothing but category of the instance. Ex. In a school student are categorized into different classes. The class is decided based on features of the student like age, result of the previous year, date of birth etc. Based on marks of the students, further they can be categorized into pass class and fail class.

6. Role of machine learning in agriculture decision making

Some of methods are already used in agriculture, those are decision making methods based on some mathematical computations. More or less those are based on basic features of machine learning. Few examples we will discuss here. A land evaluation is done based on features belong to climate and site-soil [16]. It is developed using statistics and neural network model techniques. This is bit static method not user friendly, so not used widely. Decision support system [6] is based on soil features like topography, nutrients, history of cultivation, precipitation etc. Static categorization of the feature values is done to evaluate suitability level. The drawback of this system is also a static nature. Konstantinos G Liakos has discussed in detail how machine learning played role in agriculture precision through crop management, yield prediction, disease detection, soil and water management [17]. Patricio and Rieder [18] mentioned that artificial intelligence plays important role in improving accuracy in agriculture. During 2013–2017 data captured by camera was analyzed using support vector machine classifier. Uddin, Mohammad Shorif has discussed contribution of machine learning and computer vision in agriculture [19]. Machine learning helps in decision making, for better productivity and more precise systems. In developed countries machine learning is introduced in agriculture too early for different purpose like farming prediction is done using classification [20, 21], Artificial Neural Network technique used for crop yield prediction [22]. Even for study regarding plant disease statistical and machine learning approaches has been used [23–27].

Decision making about suitable crop

As we have discussed under Section 4, we can get the feature values for a particular crop in a particular season as per the list of features identified under Section 4.1.1 and also, we can compute the level of suitability of the same crop using the methodology discussed under Section 4.2. Here, crops specific suitability is considered as output class. Suitability level is further divided into five classes [15], so classification method of machine learning is chosen. According to the discussion under Section 5, we can say that supervised technique classification is suitable for crop specific suitability analysis. We can treat the computed suitability levels based on yield as class-value for that particular record/tuple and all other feature-values as input-feature-values of the same tuple. Any classification technique, which can classify the records into more than one classes based on input features called as multi-class classification technique. Any multi-class classification technique can be used and further customized [28] to get the appropriate suitability analysis. Similar to output classes computed using Eqs. (7)–(11) input classes can be computed as below.

If input feature is categorical value then no need to compute the levels, it can be directly mapped. Example soil quality is one of the features need to consider having three different having categorical values good, moderate and average. Then it will be mapped to input levels as below:

Level1   Soil_quality = Good
Level2   Soil_quality = Moderate
Level3   Soil_quality = Average

Let us consider the input X_i is environment feature temperature at Punjab state in India. Lowest temperature is considered as 0°C and highest temperature observed is 50°C. We know in the growing phase of the crop wheat for high growth rate the favorable temperature is 20–25°C. Wheat cannot tolerate the temperature below 3.5°C so, Level5 is less than or equal to 3.5°C. It cannot tolerate the temperature above 35°C so, Level1 is above 35°C.

Thus,
Highest value is   35°C
Lowest value is   3.5°C
So, as per Eq. (1) interval is calculated as,
X_i = (35–3.5)/4 = 7.75.
X_1 is   35
X_2 is $X_1 - X_i$ = 35–7.75 = 19.5
X_3 is $X_2 - X_i$ = 27–8 = 11.75
X_4 is $X_3 - X_i$ = 21–8 = 4

To convert into input class levels, the dynamic levels will be computed. If the available dataset has tuples as shown in **Table 1**.

To simplify we will round up the values. The levels will be converted as below.

Level1 > =35°C.
Level2 < 35°C and > = 20°C.
Level3 < 20°C and > =12°C.
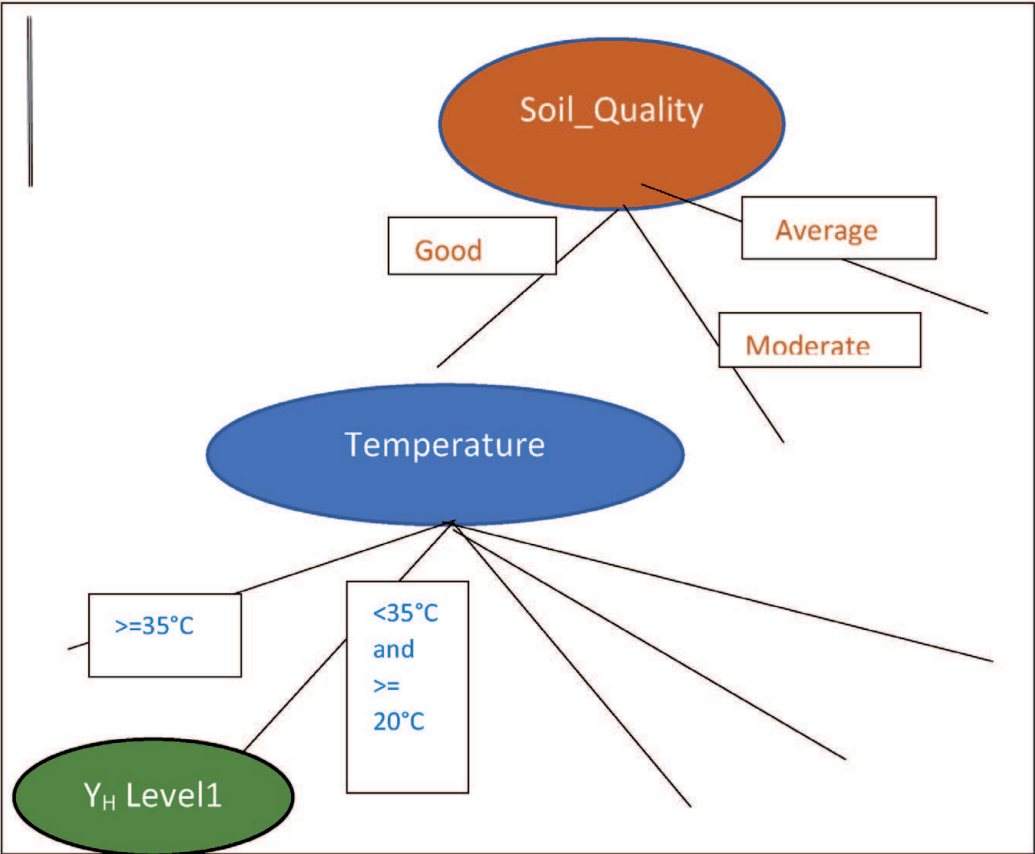Level4 < 12°C and > =4°C.
Level5 < 4°C.

Figure 1.
Decision making tree.

Now, as per tuples in database for suitability level 1 of output the mapping input level in Level2, thus for decision tree classification intermediate result based on above data belongs **Table 1**, partial decision tree will be as shown in **Figure 1**.

Based on simple dataset available as per **Table 1**, simple decision tree has been formed. If more features will be considered levels of the tree will be increased. Always the last level of the tree will output class i.e. suitability/yield level. For n number of input features tree will have n + 1 level [29].

7. Conclusion

We can conclude that machine learning can be applied in agriculture decision making. More the balanced and appropriate dataset is available better the decision can be taken. Here the machine learning approach we used, called as decision tree classification. So, we can say quality of decision tree formation is dependent on quality of input dataset. Advanced decision tree approaches work on variety of data values like categorical, constant and discrete (numerical as well as text values with some preprocessing) values. We can consider all these variety of agriculture features for processing and decision making.

Author details

Ratnmala Bhimanpallewar[1*] and M.R. Narasinga Rao[2]

1 Vishwakarma Institute of Information Technology, Pune, India

2 K.L. University, Vijayawada, India

*Address all correspondence to: ratnmalab@gmail.com

IntechOpen

References

[1] Food and Agriculture Organization of the United Nations (FAO). India Country Programming Framework. 2016; Available from: http://www.fao.org/3/a-bp575e.pdf

[2] FAO_Chapter 6 - Causes of land degradation [Internet]. United Nation: FAO; Available from: http://www.fao.org/docrep/v4360e/V4360E00.htm#Contents

[3] Timespoints TOI. Kolhapur Cancer news dated Aug 30 2018. Times of India [Internet]. 2018;1-18. Available from: http://timesofindia.indiatimes.com/articleshow/66977019.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst

[4] Ameen A, Raza S. International Journal of Advances in Scientific Research Green Revolution: A Review QR Code *Correspondence Info. Int J Adv Sci Res [Internet]. 2017;3(12):129-37. Available from: https://doi.org/10.7439/ijasr.v3i12.4410

[5] Flachs A. Encyclopedia of Food and Agricultural Ethics. Encycl Food Agric Ethics. 2016;(November).

[6] Elsheikh R, Mohamed Shariff ARB, Amiri F, Ahmad NB, Balasundram SK, Soom MAM. Agriculture Land Suitability Evaluator (ALSE): A decision and planning support tool for tropical and subtropical crops. Comput Electron Agric [Internet]. 2013;93:98-110. Available from: http://dx.doi.org/10.1016/j.compag.2013.02.003

[7] Widiatmaka, Ambarwulan W, Setiawan Y, Walter C. Assessing the suitability and availability of land for agriculture in tuban regency, East Java, Indonesia. Appl Environ Soil Sci. 2016;2016.

[8] Widiatmaka W, Ambarwulan W, Firmansyah I, Munibah K, Santoso PBK, Sudarsono S. Land suitability and dynamic system modelling to define priority areas of soybean plantation in paddy fields in Karawang, West Java. Agrivita. 2014;36(3):235-48.

[9] Karimi F, Sultana S, Shirzadi Babakan A, Royall D. Land Suitability Evaluation for Organic Agriculture of Wheat Using GIS and Multicriteria Analysis. Pap Appl Geogr [Internet]. 2018 Jul 3;4(3):326-42. Available from: https://doi.org/10.1080/23754931.2018.1448715

[10] Doraiswamy PC, Hatfield JL, Jackson TJ, Akhmedov B, Prueger J, Stern a. Crop condition and yield simulations using Landsat and MODIS. Remote Sens Environ. 2004;92:548-59.

[11] Chandrasekaran B, Annadurai K, Somasundaram E. A Text book of Agranomy. 2010. 1-856 p.

[12] Services IA-I and. No Title [Internet]. Nagarjuna Education Trust, Hyderabad, India. 2020. Available from: http://www.ikisan.com

[13] Agriculture University. Krishi Darshani 2008. Pune; 2008.

[14] Ratnmala Bhimanpallewar, M. R. Narasinagrao, A prototype model for c ontinuous agriculture field monitoring and assessment, International Journal of Engineering & Technology, 2018; 7 (2.7): 179-182.

[15] FAO. Chapter 3 : Land suitability classifications. A Framew L Eval Soils Bull 32 [Internet]. 1976;1-9. Available from: http://www.fao.org/3/X5310E/x5310e00.htm

[16] D. De La Rosa, F. Mayol, E. Diaz-Pereira, *M. Fernandez*, and D. De La Rosa, A land evaluation decision support system (MicroLEIS DSS) for agricultural soil protection: With special

reference to the Mediterranean region, *Environ. Model. Softw.*, 2004; 19(10):929-942.

[17] Konstantinos G. Liakos, Machine Learning in Agriculture: A Review, Sensors, 2018; 18(8):2674:1-29.

[18] Jha K, Doshi A, Patel P, Shah M, Artificial Intelligence in Agriculture, 2019, 2: 1-12.

[19] Uddin, Mohammad Shorif, Computer Vision and Machine Learning in Agriculture, ISBN 978-981-336-424-0.

[20] Washington Okori, Joseph Obua,Machine Learning Classification Technique for Famine Prediction, Proceedings of the World Congress on Engineering, WCE 2011, London, U.K, 2011 Jul; II

[21] S.R.Rajeswari , Parth Khunteta, Subham Kumar,Amrit Raj Singh,Vaibhav Pandey, Smart Farming Prediction Using Machine Learning, International Journal of Innovative Technology and Exploring Engineering ISSN: 2278-3075, 2019 May; 8(7): 190-194

[22] Miss.Snehal, S.Dahikar, Dr.Sandeep V.Rode, "Agricultural Crop Yield Prediction Using Artificial Neural Network Approach". International Journal of Innovative Research in Electrical, Electronic, Instrumentation and Control Engineering, Vol. 2, Issue 1, January 2014

[23] Gutierrez D. D., Machine Learning and Data Science: An Introduction to Statistical Learning Methods with R. Basking Ridge, NJ: Technics Publications, 2015

[24] Mehra, Lucky K. et al., Predicting Pre-Planting Risk of Stagonospora Nodorum Blotch in Winter Wheat Using Machine Learning Models, Frontiers in Plant Science, 2016 Apr; PMC. Web. 16, 7 (2016): 390

[25] Kim KS, Beresford RM, Development of a disease risk prediction model for downy mildew (Peronospora sparsa) in boysenberry., Walter M Phytopathology. 2014 Jan; 104(1):50-6.

[26] Rumpf, T., A.-K. Mahlein, U. Steiner, E.-C. Oerke, H.-W. Dehne, and L. Plümer., Early Detection and Classification of Plant Diseases with Support Vector Machines Based on Hyperspectral Reflectance, Computers and Electronics in Agriculture 74.1 (2010): 91-99. Web.

[27] M.P.Raj, P,R,Swaminarayan, J.R.Saini, D.K.Parmar, Application Of Pattern Recognition Algorithm In Agriculture : A Review, Int.J.Advanced Networking and Application, 2015; 6(5).

[28] Ratnmala Bhimanpallewar, M. R. Narasinagrao, A Machine Learning Approach to Assess Crop Specific Suitability for Small/Marginal Scale Croplands, International Journal of Applied Engineering Research, 2017 Nov;12(23):13966-73.

[29] Ratnmala Bhimanpallewar, M. R. Narasinagrao, Evaluating the Influence of Soil and Environmental Parameters in Terms of Crop Suitability using Machine Learning, Indian Journal of Agricultural Research, 2021; DOI:10.18805/IJARe.A-4942