Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Predictive soil-crop suitability pattern extraction using machine learning algorithms

Kristine T. Soberano [1], Jeffric S. Pisueña [1,*], Shara Mae R. Tee [2], Jan Carlo T. Arroyo [3,4], Allemar Jhone P. Delima [3,4]

[1]*Faculty of Information Technology, Northern Negros State College of Science and Technology, Sagay, Philippines*
[2]*Faculty of Information Technology, Central Philippine State University, Kabankalan, Philippines*
[3]*College of Information and Computing Studies, Northern Iloilo State University, Estancia, Iloilo, Philippines*
[4]*College of Computing Education, University of Mindanao, Davao City, Davao del Sur, Philippines*

A R T I C L E I N F O

Article history:
Received 7 October 2022
Received in revised form
23 February 2023
Accepted 4 April 2023

Keywords:
Data mining
Machine learning algorithms
Pattern extraction
Soil-crop suitability

A B S T R A C T

Machine learning has experienced notable advancements in recent times. Furthermore, this field facilitates the automation of human evaluation and processing, leading to a reduced demand for manual labor. This research paper employs data mining techniques and Knowledge Discovery in Databases (KDD) to conduct an evaluation and classification of various algorithms for pattern extraction and soil suitability prediction. The study utilizes experimental data, data transformation, and pattern extraction techniques on diverse soil samples obtained from different regions of Negros Occidental, Philippines. Specifically, the Naive Bayes, Deep Learning, Decision Tree, and Random Forest algorithms are selected for the classification and prediction of soil suitability based on the available datasets. The assessment of soil-crop suitability is based on data sourced from the Philippine Rice Research Institute, considering 14 parameters including inherent fertility, soil pH, organic matter, phosphorus, potassium, nutrient retention (CEC), base saturation, salinity hazard, water retention, drainage, permeability, stoniness, root depth, and erosion. The findings indicate that the Random Forest algorithm achieved the highest accuracy rate at 94.6% and the lowest classification error rate at 5.4%, suggesting a high level of confidence in the model's predictions. The model's predictions reveal that most soil samples in the area are only marginally suitable for banana, maize, and papaya crops. Furthermore, the study demonstrates that the majority of soil samples have a low fertility rating, which significantly impacts crop suitability. The information obtained from this study can serve as a basis for local farmers to develop improved soil management programs aimed at ensuring more productive soil. Simultaneously, it can contribute to active soil protection initiatives addressing issues such as acidity and salinity in Negros Occidental, Philippines.

1. Introduction

Crop suitability indices are innovative management techniques that can be used to pinpoint the best agricultural production regions at the farm, tribal, or regional levels. They are a crucial step in achieving sustainable intensification (Smith et al., 2022). For many farmers, in particular, soil-crop suitability is a requirement for agricultural improvement (Bhimanpallewar and Narasingarao, 2022). The origin, features, and significance of the soil in relation to crop growth are known through the classification of soil and crop suitability. The main elements considered when determining whether a plot of land is suitable for growing crops include the land-use-land-cover (LULC), TWI, altitude, level, precipitation, soil properties, landform, groundwater, and land surface temperature (Ramu et al., 2022).

Problems related to organic matter, salinization, soil compaction, alkalinization, and carbonates restrict the suitability of crops (AbdelRahman and Arafat, 2020). On the other hand, Machine Learning is a computer science field where new developments have recently been developed, which also helps to automate assessment and processing carried out by

* Corresponding Author.
Email Address: jpisuena@nonescost.edu.ph (J. S. Pisueña)
https://doi.org/10.21833/ijaas.2023.06.002
Corresponding author's ORCID profile:
https://orcid.org/0000-0003-1372-035X

mankind so that human manual power is reduced. Artificial intelligence (AI) known as machine learning enables computers to learn without being explicitly programmed (Martis et al., 2022). In this paper, the researchers aim to utilize the Knowledge Discovery in Database (KDD) to perform an evaluation and classification of the different algorithms for pattern extraction and prediction of soil suitability. Knowledge discovery in databases (KDD) is the method of acquiring knowledge from a collection of data. This popular data mining technique is a procedure that includes data selection, cleansing, and incorporation prior knowledge of data sets and accurate interpretation solutions based on the measured outcome (Hlaing and Thaw, 2019).

This paper also employs experimental data, data transformation, and pattern extraction techniques on varied soil samples from different parts of Negros Occidental, Philippines. The Naive Bayes, Deep Learning, Decision Tree, and Random Forest are the chosen algorithms that will derive new data sets. With the derived data sets and model prediction framework, this study would be a great benefit to farmers to plant crops that are suitable for soil qualities in their areas.

2. Related works

According to Agarwal and Tarar (2021), plant weeds are defined as plant species that are present at the incorrect time and location. Selecting a crop that is not suitable for a particular type of location is one of the causes of low crop yield production. Since the sample dataset includes growth information, it will assist in deciding on the best type of soil for seeds. The extracted features are compared to the characteristics of the sample dataset. In this approach, prediction is only made following the expansion of results in a decline in the crop's growth quality. The seeds are categorized by a machine algorithm based on their growth and forecasting of agricultural diseases. Training the dataset accomplishes this by comparing the characteristics of newly harvested seeds to those of sample seeds and forecasting the development of illnesses in the crop.

Meanwhile, John et al. (2020) utilized AI algorithms will lessen the farmers' difficulties in obtaining losses in their farms as a result of their ignorance of how and what plants to grow in different types of soil and climates. The algorithm identifies the perfect crops that should be planted on the least expensive land out of all the crops that are currently available after looking at the prediction parameters.

The study of Kalichkin et al. (2021) showed that when building the models, decision trees were developed using varied techniques to forecast the spring wheat output based on qualitative characteristics. The manner of soil cultivation, the placement of the crop following steam, and agro meteorological resources were among the qualitative data that were taken into account; like the sum of active air, temperature, and precipitation. The system's generality and adaptability to various natural and agricultural situations, the fewest possible input parameters (public data) were used.

The study of Rahman et al. (2018) led to the creation of a model for forecasting soil series and offering acceptable crop yield suggestions for a certain type of soil. They discovered that SVM was 94.95% accurate and can provide the best level of soil categorization among other predicting models.

3. Methodology

To perform the evaluation and classification of the different algorithms for pattern extraction and prediction of the soil suitability results the researchers used the Knowledge Discovery in Database (KDD) methodology. Fig. 1 shows the step-by-step process in KDD.

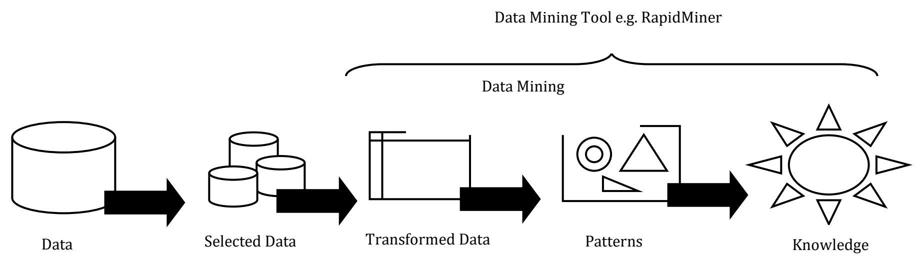

Fig. 1: Knowledge discovery in databases process

3.1. Experimental data

The dataset used in this study was obtained from the Philippine Rice Research Institute and soil data from local agriculture office of the different cities and municipalities which consists of soil fertility indicators and physical soil qualities from eighteen (18) different places in Negros Occidental namely: Bago, Bantay, Batuan, Bolinao, Cadiz, Faraon, Guimbalon, Isabela, La Castellana, Luisiana, Manapla, Obando, Pulupandan, Silay, San Manuel, Tupi, Umingan, and Victorias with a total of 111 soil samples. The following soil fertility indicators serve as the source for the metrics that are used: Phosphorus; pH; organic matter; base saturation; inherent fertility; hazard for salinity, potassium; and nutrient retention (CEC). Physical and soil qualities were also included in the parameters, namely: water retention, drainage, permeability, stoniness, root depth, and erosion.

As bases for the soil-crop suitability, soil fertility and qualities were matched based on three (3) crops namely: Banana, Maize, and Papaya. Suitability ratings are based on the data from the Philippine Rice Research Institute and local agriculture offices of the different locations. The suitability rating is used as a label for processing data using a data mining algorithm for pattern extraction and prediction as shown in Table 1.

Table 1: Suitability ratings

Label	Description
S1	Highly suitable
S2	Moderately suitable
S3	Marginally suitable
N1	Currently not suitable
N2	Permanently not suitable

3.2. Data transformation

In transforming the data, a series of steps were followed. This started with data selection of soil fertility indicators and physical soil quality parameters. Data cleaning was also done to ensure no missing values and incorrect data were included in the selected data. Out of 118 soil samples, 111 soil samples was selected for data analysis. The dataset has 16 parameters in total, of which the crop suitability parameter was the target variable for prediction. Data was encoded using MS Excel application and then saved as a Comma Separated Values (CSV) file. The description of the dataset is shown in Table 2.

Table 2: Dataset description

Parameters	Description
Location	The place where the soil samples were collected
Crop Suitability Rating	The rating based on Simplified Keys to Soil Series
Inherent Fertility	1-none; 2-very low; 3-low; 4-moderate; 5-high; 6-very high
Soil pH	1-acidic; 2-slightly acid; 3-neutral; 4-slightly alkaline; 5-alkaline
Organic Matter	1-none; 2-very low; 3-low; 4-moderate; 5-high; 6-very high
Phosphorus (P)	1-none; 2-very low; 3-low; 4-moderate; 5-high; 6-very high
Potassium (K)	1-none; 2-very low; 3-low; 4-moderate; 5-high; 6-very high
CEC (Nutrient Retention)	1-none; 2-very low; 3-low; 4-moderate; 5-high; 6-very high
Base Saturation	1-none; 2-very low; 3-low; 4-moderate; 5-high; 6-very high
Salinity Hazard	1-none; 2-very low; 3-low; 4-moderate; 5-high; 6-very high
Water retention	1-very low; 2-low, 3-moderate; 4-high; 5-very high
Drainage	1-very slow; 2 moderate; 3 good; 4 excessive
Permeability	1-very slow; 2-slow; 3-moderate; 4-rapid
Stoniness	1-none; 2-common; 3-many; 4-abundant; 5-pebbles and outcrops; 6-stones and boulders; 7-boulder outcrops
Root Dept	1-very shallow; 2-shallow; 3-moderate; 4-deep; 5-very deep
Erosion	1-none; 2-moderate; 3-severe

3.3. Pattern extraction

To illustrate the process of pattern extraction and prediction, the researchers adopted the model prediction framework from S.Y. Muhammad et al. (2015) as shown in Fig. 2 (Muhammad et al., 2015). This study used pre-processed data needed. This study used different algorithms as classifiers for model prediction and pattern extractions namely the Naïve Bayes, Deep Learning (H2O), Decision Tree, and Random Forest. After running the different models, the accuracy and performance of each algorithm were then evaluated and interpreted.

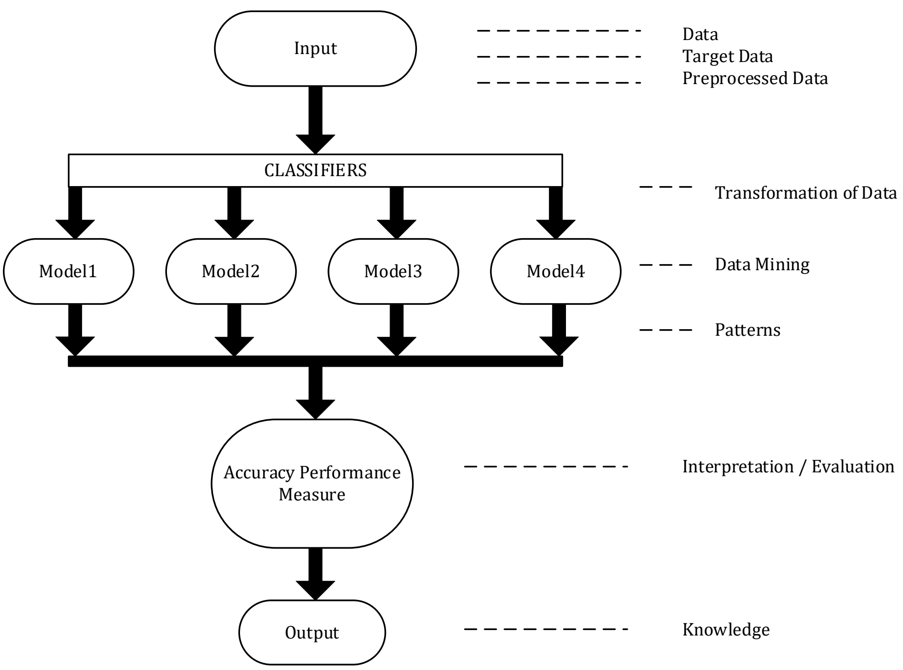

Fig. 2: Model prediction framework (Muhammad et al., 2015)

4. Results and discussion

The soil fertility and soil physical qualities data were collected from 18 different locations in Negros Occidental. The locations were the following: L1 (Bago), L2 (Bantay), L3 (Batuan), L4 (Bolinao), L5 (Cadiz), L6 (Faraon), L7 (Guimbalaon), L8 (Isabela), L9 (La Castellana), L10 (Luisiana), L11 (Manapla), L12 (Obando), L13 (Pulupandan), L14 (Silay), L15 (San Manuel), L16 (Tupi), L17 (Umingan), L18 (Victorias).

4.1. Exploratory data analysis

Inherent fertility indicates the capacity of soil to retain and release nutrients for the plant. It is noted that majority of the soil samples has low inherent fertility level (Guimbalon, Obando, and Victorias) followed by moderate and high as shown in Fig. 3. No soil sample with none, very low and very high level of inherent fertility.

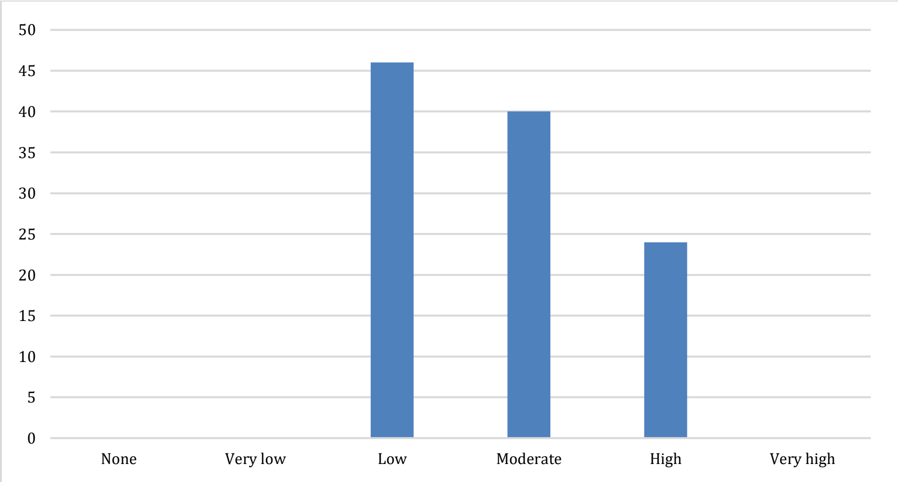

Fig. 3: Inherent fertility distribution

Soil pH or Soil Acidity lowers the availability of certain nutrients for plants, like phosphorus and molybdenum, while raising the supply of other elements to unsafe levels (Ni et al., 2018). Fig. 4 shows that out of 18 locations, 7 locations (81 soil samples) were found to have acidic soil based on soil samples namely: Batuan (L3), La Castellana (L9), Luisiana (L10), Manapla (L11), Silay (L14), Tupi (L16) and Victorias (L18).

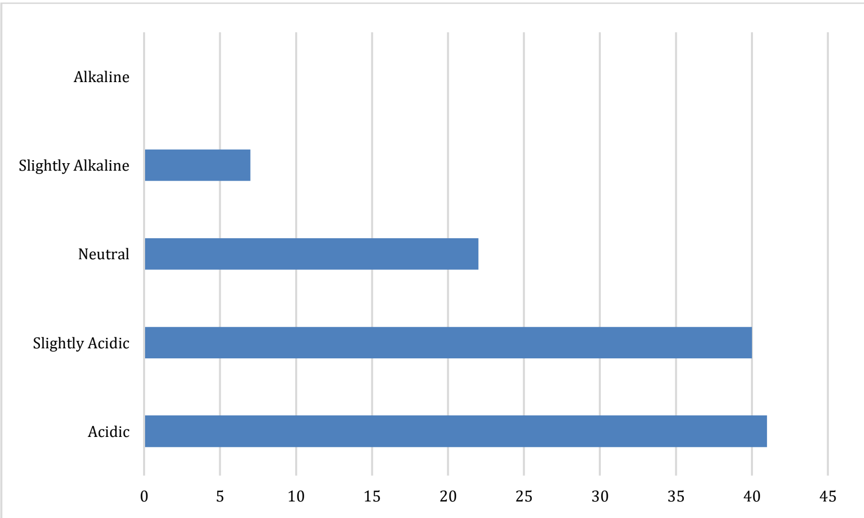

Fig. 4: Soil acidity distribution

Having high level of Phosphorus may tend to cause harm to the overall plant health since this may cause deficiencies in other minerals in the soil like zing and iron. It is well noted that several soil samples (47) contain a high level of phosphorus as shown in Fig. 5. These samples are from Bago (L1), Bantay (L2), Bolinao (L4), Isabela (L8), Silay (L14), San Manuel (L15), Tupi (L16), and Umingan (L17).

Soil Salinity greatly affects plant growth and production (Hafez et al., 2021). Salinity affects plants by interfering the nitrogen uptake thus reducing growth and reproduction. In Fig. 6, twelve (12) soil samples from Obando (L12) was identified with a moderate level of Salinity Hazard which may be considered a limiting factor in crop growth and production.

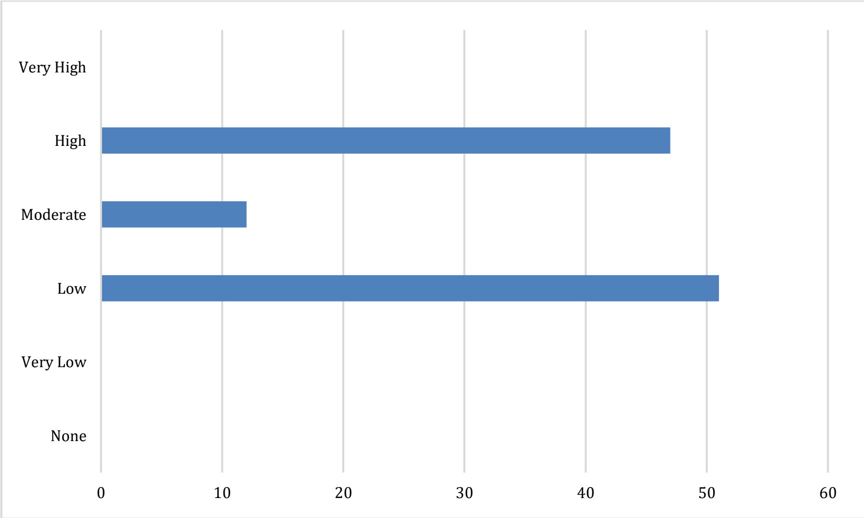

Fig. 5: Phosphorus distribution

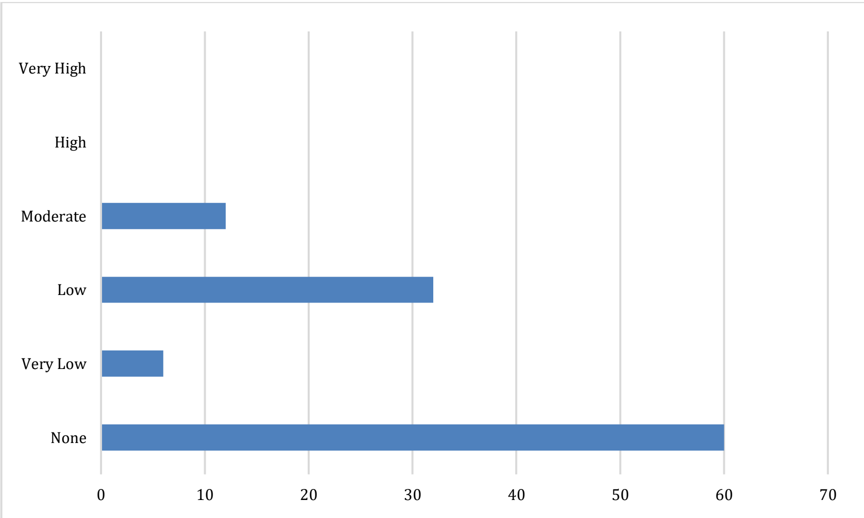

Fig. 6: Salinity hazard distribution

Based on the result as shown in Fig. 7, the marginally suitable (S3) rating is 78.18%, the currently not suitable (N1) rating is 16.36%, and the permanently not suitable (N2) rating is 5.45%. There was no sample with a rating of moderately suitable (S2) and highly suitable (S1) for the banana crop. This demonstrates that the majority of soil samples are adequate for the banana crop, but some may need appropriate soil management that concentrates on improving soil fertility, lowering soil acidity, and reducing salinity risks.

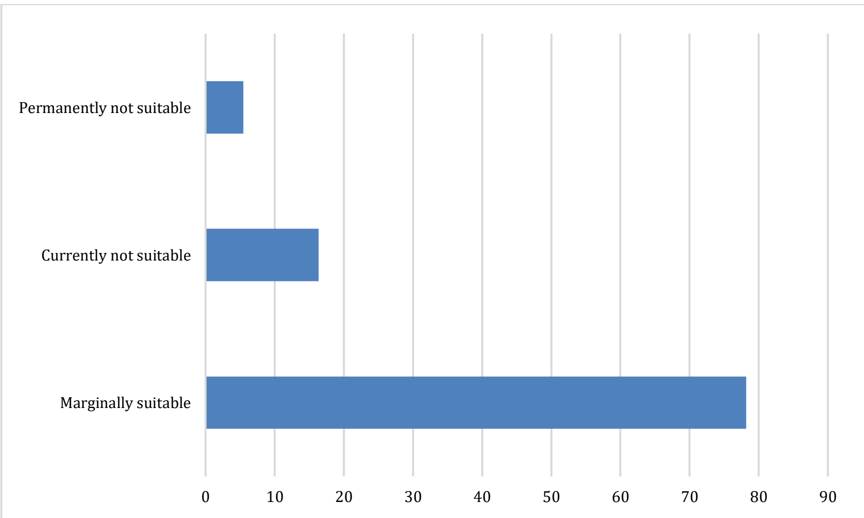

Fig. 7: Summary of Banana suitability rating

The result shows (Fig. 8) that the rating of marginally suitable is 89.1%, currently not suitable is 5.45%, and permanently not suitable is also 5.45%. Still, there was no sample with ratings of moderately suitable (S2) and highly suitable (S1) for the maize crop. The maize crop is generally suitable in most sampled soil, though the same with the banana crop, soil management will play a vital role in ensuring better suitability.

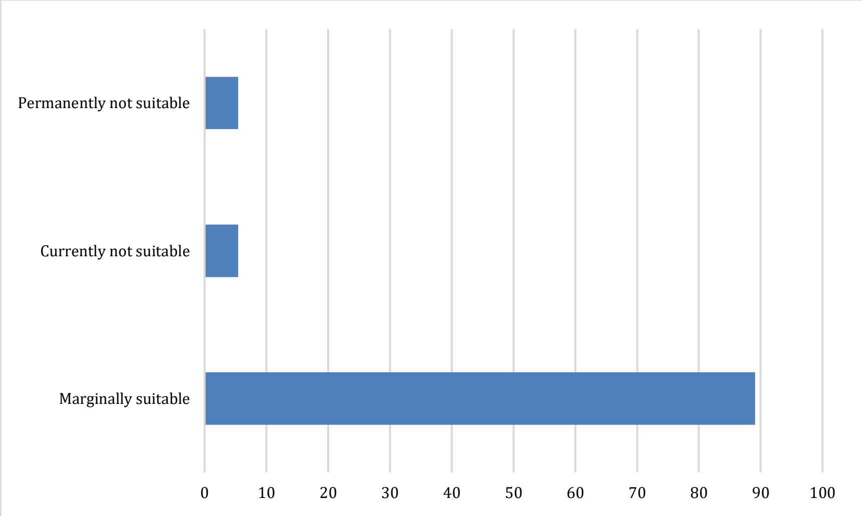

Fig. 8: Summary of maize suitability rating

Generally, most of the soil samples were rated as currently not suitable with 38.18%, followed by marginally suitable with 34.55%, moderately suitable with 16.36%, and permanently not suitable with 10.91%. No sample with a highly suitable (S1) rating as shown in Fig. 9. This shows that most sampled soil were currently not suitable for papaya crop but was followed by a marginally suitable rating. Inherent fertility, water retention, and, CEC (nutrient retention) has the highest weight in the ranking.

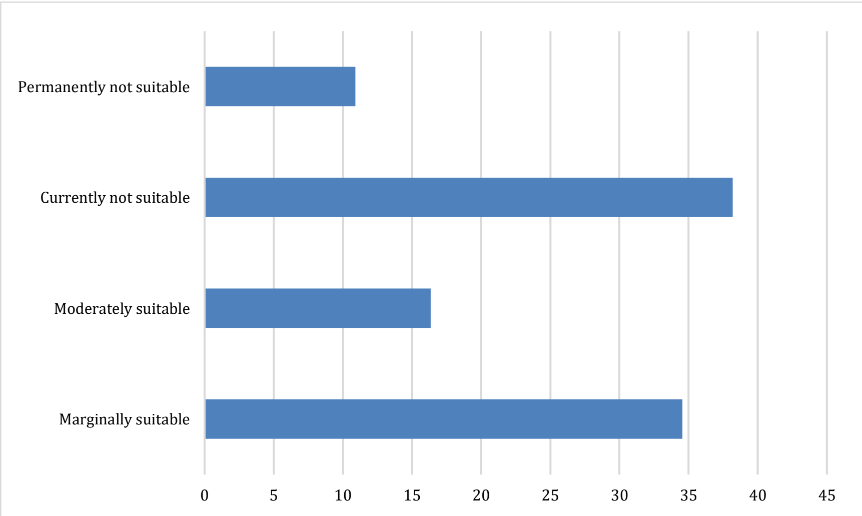

Fig. 9: Summary of papaya suitability rating

4.2. Algorithm results

Among the different algorithms, the random forest has the highest accuracy (average of 94.6%)-having the lowest classification error when used on three (3) different crops as shown on Table 3. This is followed by deep learning (H2O) (average of 89.4%), decision trees, and naive Bayes. The accuracy of the classification signifies how correct the prediction of each algorithm is in determining crop-soil suitability.

Table 3: Algorithm performance

Algorithm	Accuracy (banana suitability)	Accuracy (maize suitability)	Accuracy (papaya suitability)	Average accuracy
Naive Bayes	77.1%	90.5%	45.7%	71.1
Deep learning (H2O)	96.7%	90.5%	81.0%	89.4
Decision tree	93.3%	90.5%	78.1%	87.3
Random forest	96.7%	90.5%	96.7%	94.6

The confusion matrix and models shown in the following tables were based on Random Forest given that this algorithm showed the highest accuracy and precision.

Table 4 presents the precision of the prediction. The class label is the suitability rating based on Table 1. Table 1 shows that marginally suitable (S3) got 100%, and currently not suitable (N1) got 83.33% which were all acceptable predictions. Permanently not suitable (N2) got 0% due to not balanced data for each label.

Table 4: Confusion matrix-banana suitability

	True S3	True N1	True N2	Class precision
pred. S3	24	0	0	100.0%
pred. N1	0	5	1	83.33%
pred. N2	0	0	0	0.0%
class recall	100%	100%	0.00%	

Table 5 shows that marginally suitable (S3) got 90.32% accuracy which is an acceptable prediction. N1 and N2 has no result, still due to imbalanced data for each label and since most data is leaning on S3-marginally suitable.

Table 5: Confusion matrix-maize suitability

	True N1	True S3	True N2	Class precision
pred. N1	0	0	0	0.0%
pred. S3	2	28	1	90.32%
pred. N2	0	0	0	0.0%
Class recall	0.0%	100%	0.0%	

Table 6 shows that the rating currently not suitable (N1) got 100%, marginally suitable (S3) got 100%, permanently not suitable (N2) got 100% and moderately suitable got 75%, which are all good and acceptable predictions.

Table 6: Confusion matrix-papaya suitability

	True N1	True S3	True N2	True S2	Class precision
pred. N1	15	0	0	0	100%
pred. S3	0	9	0	0	100%
pred. N2	0	0	2	0	100%
pred. S2	0	1	0	3	75%
Class recall	100%	90%	100%	100%	

The overall result of prediction accuracy is generally acceptable thus the model is considered acceptable. The developed model is shown in Table 7, Table 8, and Table 9. As shown in Table 7, Salinity Hazard (SH) has a higher weight in determining the suitability of the banana crop, this is followed by P (phosphorous) and the inherent fertility (IF) of soil.

For maize crop suitability, stoniness, phosphorous, permeability, root deep, drainage, and CEC play a vital role (Table 8). Having a low CEC greatly contributes to the soil's suitability for maize crops. The model for Papaya crop suitability shows that salinity hazard (SH), pH, drainage, base saturation (BS), permeability (PERM), CEC, and stoniness (S) of soil greatly affects its suitability.

5. Conclusion

The utilization of data mining techniques for extracting concealed knowledge from a dataset is widely employed to harness the derived information for specific fields of study. In this particular investigation, data mining techniques are employed to assess the suitability of soil for different crops based on soil samples obtained from Negros Occidental. The researchers considered 14 parameters, namely inherent fertility, soil pH, organic matter, phosphorus, potassium, nutrient retention (CEC), base saturation, salinity hazard, water retention, drainage, permeability, stoniness, root depth, and erosion, to analyze soil-crop suitability. The soil-crop suitability ratings were derived from data provided by the Philippine Rice Research Institute.

For pattern extraction and prediction purposes, the researchers employed four models. Among these models, the Random Forest algorithm achieved the highest accuracy (94.6%) and the lowest classification error (5.4%), indicating a high level of confidence in the model's predictive capabilities. The results revealed that the majority of the soil in the area is marginally suitable for banana cultivation (with 100% prediction accuracy for S3 and 83.33% prediction accuracy for N1), maize cultivation (with 90.32% prediction accuracy for S3), and papaya cultivation (with 100% prediction accuracy for N1, S3, N2, and 75% prediction accuracy for S2). However, it should be noted that the papaya crop exhibited a relatively higher proportion of unsuitability.

The study also uncovered that the majority of the soil samples had low fertility ratings, significantly impacting crop suitability. This information can serve as a foundation for the development of enhanced soil management programs aimed at ensuring increased soil productivity and active soil protection in terms of acidity and salinity in Negros Occidental, Philippines.

Table 7: Developed model for banana crop suitability

If	Then
SH>3.5	Currently not suitable
SH<=3.5 and D>1.5 and PH>1.5	Marginally suitable
SH<=3.5 and D<=1.5 and P<=4 and PERM>2.5	Marginally suitable
SH<=3.5 and D<=1.5 and P<=4 and PERM>2.5	Permanently not suitable
PH<=3.5 and PERM>2.5 and P>4	Marginally suitable
PH<=3.5 and PERM >2.5 and P<=4 and SH<=3.5 and D>2.5	Currently not suitable
PH<=3.5 and PERM>2.5 and P<=4 and SH<=3.5 and D<=2.5	Marginally suitable
K>3.5	Marginally suitable
K<=3.5 and SH<=3.5 and SH>2.5 and PH>1.5	Permanently not suitable
K<=3.5 and SH<=3.5 and SH>2.5 PH<=1.5	Marginally suitable
S>1.5 and PERM>2.5	Marginally suitable
S<=1.5 and P<=4.5 and WR>1.5	Currently not suitable
S<=1.5 and P<=4.5 and WR<=1.5	Marginally suitable
IF>3.5 and D>1.5 and BS>3.5	Marginally suitable
IF>3.5 and D<=1.5 and P>4	Marginally suitable
IF>3.5 and D<=1.5 and P<=4	Permanently not suitable
IF<=3.5 and BS>4.5	Currently not suitable
IF<=3.5 and BS<=4.5 and S>1.5 and D>1.5	Marginally suitable
IF<=3.5 and BS<=4.5 and S>1.5 and D<=1.5	Permanently not suitable
IF<=3.5 and BS<=4.5 and S<=1.5 and P<=3.5	Marginally suitable

Table 8: Developed model for maize crop suitability

If	Then
CEC>4.5	Marginally suitable
CEC<=4.5 and PERM>3.5 and D>3.5	Currently not suitable
CEC<=4.5 and PERM>3.5 and D<=3.5	Marginally suitable
CEC<=4.5 and PERM<=3.5 AND PERM>2.5	Marginally suitable
CEC<=4.5 and PERM<=3.5 and PERM>2.5 and S>3	Permanently not suitable
S>7.5	Currently not suitable
S<=7.5 and WR>3.5 and OM>3.5	Marginally suitable
S<=7.5 and WR>3.5 and OM<=3.5	Permanently not suitable
S <=7.5 and WR<=3.5	Marginally suitable
P > 3.5 and RD>2.5 and BS>4	Marginally suitable
P>3.5 and RD>2.5 and BS<=4	Currently not suitable
P<=3.5 and S>3.5 and PERM>2.5	Marginally suitable
P<=3.5 and S>3.5 and PERM<=2.5	Permanently not suitable
P<=3.5 and S<=3.5	Marginally suitable
WR>2.5 and RD>1.5	Marginally suitable
WR>2.5 and RD<=1.5	Permanently not suitable
WR<=2.5 and K<=3.5 and D>3.5	Currently not suitable
WR <=2.5 and K<=3.5 and D<=3.5	Marginally suitable
D<=1.5 and K<=3.5 and pH>2.5	Permanently not suitable
D<=1.5 and K<=3.5 and pH<=2.5	Marginally suitable

Table 9: Developed model for papaya crop suitability

IF	Then
PH>1.5	Currently not suitable
K<=3.5 and PERM>1.5 and CEC>4.5	Moderately suitable
S<=7.5 and OM>2.5 and E>2.5 and S>5.5 and K>3.5	Marginally suitable
S<=7.5 and OM>2.5 and E<=2.5 and OM>3.5 and SH>2.5	Currently not suitable
S<=7.5 and OM>2.5 and E<=2.5 and OM>3.5 and SH<=2.5 and WR>2.5 and P>4.5 and PH>2.5	Permanently not suitable
S<=7.5 and OM>2.5 and E<=2.5 and OM>3.5 and SH<=2.5 and WR>2.5 and P>4.5 and PH<=2.5	Moderately suitable
CEC>4.5 and S<=7.5 and SH>3.5	Currently not suitable
CEC>4.5 and S<=7.5 and SH<=3.5 and D>2.5 and BS>2.5 and PERM>3.5	Marginally suitable
CEC>4.5 AND S<=7.5 and SH<=3.5 and D>2.5 and BS>2.5 and PERM<=3.5	Currently not suitable
CEC>4.5 and K>3.5 and SH>2	Currently not suitable
CEC>4.5 and K>3.5 and SH<=2	Marginally suitable
CEC>4.5 and K><=3.5	Moderately suitable
CEC<=4.5 and PERM>1.5 and OM>3.5 and PERM>2.5 and E>2.5 and WR>3.5	Currently not suitable
CEC<=4.5 and PERM>1.5 and OM>3.5 and PERM>2.5 and E>2.5 and WR<=3.5	Marginally suitable
CEC<=4.5 and PERM>1.5 and OM>3.5 and PERM>2.5 and E<=2.5 and D>3.5	Moderately suitable
CEC<=4.5 and PERM>1.5 and OM>3.5 and PERM>2.5 and E<=2.5 and D<=3.5 and SH>2	Currently not suitable
CEC<=4.5 and PERM>1.5 and OM>3.5 and PERM>2.5 and E<=2.5 and D<=3.5 and SH<=2 and K<=3.5	Currently not suitable
SH<=2.5 and WR>2.5 and K>3.5 and IF>4.5	Marginally suitable
SH<=2.5 and WR>2.5 and K>3.5 and IF<=4.5	Permanently not suitable
SH<=2.5 and WR>2.5 and K>3.5 and BS >3.5 and CEC>4.5	Moderately suitable

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

AbdelRahman MA and Arafat SM (2020). An approach of agricultural courses for soil conservation based on crop soil suitability using geomatics. Earth Systems and Environment, 4: 273-285. https://doi.org/10.1007/s41748-020-00145-x

Agarwal S and Tarar S (2021). A hybrid approach for crop yield prediction using machine learning and deep learning algorithms. Journal of Physics Conference Series, 1714(1): 012012. https://doi.org/10.1088/1742-6596/1714/1/012012

Bhimanpallewar RN and Narasingarao MR (2022). Evaluating the influence of soil and environmental parameters in terms of crop suitability using machine learning. Indian Journal of Agricultural Research, 56(2): 208-213. https://doi.org/10.18805/IJARe.A-4942

Hafez EM, Osman HS, Gowayed SM, Okasha SA, Omara AED, Sami R, and Abd El-Razek UA (2021). Minimizing the adversely impacts of water deficit and soil salinity on maize growth and productivity in response to the application of plant growth-promoting rhizobacteria and silica nanoparticles. Agronomy, 11(4): 676. https://doi.org/10.3390/agronomy11040676

Hlaing KS and Thaw YMKK (2019). Applications, techniques and trends of data mining and knowledge discovery database. International Journal of Trend in Scientific Research and Development, 3(5): 1604-1606.

John K, Abraham Isong I, Michael Kebonye N, Okon Ayito E, Chapman Agyeman P, and Marcus Afu S (2020). Using machine learning algorithms to estimate soil organic carbon variability with environmental variables and soil nutrient indicators in an alluvial soil. Land, 9(12): 487. https://doi.org/10.3390/land9120487

Kalichkin VK, Alsova OK, and Maksimovich KY (2021). Application of the decision tree method for predicting the yield of spring wheat. In the IOP Conference Series: Earth and Environmental Science, IOP Publishing, Surakarta, Indonesia: 032042. https://doi.org/10.1088/1755-1315/839/3/032042

Martis JE, Sannidhan MS, and Sudeepa KB (2022). A farmer-friendly connected IoT platform for predicting crop suitability based on farmland assessment. In the Internet of Things and Analytics for Agriculture, Springer, Singapore, Singapore: 247-272. https://doi.org/10.1007/978-981-16-6210-2_12

Muhammad SY, Makhtar M, Rozaimee A, Aziz AA, and Jamal AA (2015). Classification model for water quality using machine learning techniques. International Journal of Software Engineering and Its Applications, 9(6): 45-52. https://doi.org/10.14257/ijseia.2015.9.6.05

Ni K, Shi YZ, Yi XY, Zhang QF, Fang L, Ma LF, and Ruan J (2018). Effects of long-term nitrogen application on soil acidification and solution chemistry of a tea plantation in China. Agriculture, Ecosystems and Environment, 252: 74-82. https://doi.org/10.1016/j.agee.2017.10.004

Rahman SAZ, Mitra KC, and Islam SM (2018). Soil classification using machine learning methods and crop suggestion based on soil series. In the 21st International Conference of Computer and Information Technology, IEEE, Dhaka, Bangladesh: 1-4. https://doi.org/10.1109/ICCITECHN.2018.8631943

Ramu P, Sai Santosh B, and Chalapathi K (2022). Crop-land suitability analysis using geographic information system and remote sensing. Progress in Agricultural Engineering Sciences, 18(1): 77-94. https://doi.org/10.1556/446.2022.00050

Smith HW, Ashworth AJ, and Owens PR (2022). GIS-based evaluation of soil suitability for optimized production on US tribal lands. Agriculture, 12(9): 1307. https://doi.org/10.3390/agriculture12091307