



Machine learning and Multi-Layer Perceptron (MLP) Modeling of Zea Mays L. Responses to Tillage and Soil Amendments

Iman Ahmadi*

Department of Agronomy and plant Breeding, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

Received 05 November 2023; Accepted 05 Jun 2024

Abstract

To model the effect of two tillage operations (i.e. conventional and minimum tillage) and seven soil amendments (i.e. C, F, RF, RFM, RTiP, RML, RTiM) on the responses of Zea Mays L. (i.e. corn and stover yields, plant height at 6th and 10th leaf phases, and relative chlorophyll content of the crop leaves at 6th and 10th leaf phases), two-class and four-class classification modeling using the machine learning and multi-layer perceptron principles was performed. To examine the effect of different algorithms considered in the models (i.e. Decision Tree Classifier, Support Vector Machine (SVM) Classifier, K-Nearest Neighbors (KNN) Classifier, and Naive Bayes Classifier as standard Machine Learning (ML) algorithms, and Multi-Layer Perceptron (MLP) Classifier as a Deep Learning (DL) algorithm) on the model performance, classification accuracy and confusion matrix, as well as precision, recall and F1 score indicators were used as the model evaluation metrics. According to the results of this study, among the standard ML algorithms considered herein, application of the SVM classifying algorithm led to relatively higher modeling accuracies; therefore, the SVM algorithm was selected as the most appropriate ML algorithm in this research. Furthermore, when the SVM algorithm was used to classify different corn yield values and the number of classes increased from 2 to 4, the accuracy of the model reduced from 0.97 to 0.82; therefore, there is a trade-off between the number of classes and the accuracy of the model. Moreover, similarity between the result of the model developed herein regarding the effect of tillage type and soil amendments on corn yield classes and the ANOVA result of the other study conducted on similar dataset, acted as cross checking for the appropriateness of the model developed in this study. Finally, application of the MLP algorithm to classify each of the dependent variables considered herein, resulted in higher accuracies compared to the accuracies of the other standard ML algorithms.

Keywords: Classification; Machine learning; Model accuracy; Multi-Layer Perceptron; Soil amendments; Tillage

1. Introduction

Agricultural productivity can be affected, among other things, by the amount and type of soil amendments given to the soil, and type of tillage operations performed ([1], [2]). To demystify the underlying law that governs the level of soil productivity caused by the amount and type of soil amendments used and the type of tillage operations performed, researchers studied the effect of applying organic resources, mineral fertilizers as well as the combination of both on soil productivity ([3], [4], [5]). Effect of the type of tillage operations carried out on soil productivity has also been studied by some researchers ([6], [7], [8]). In these studies, procedures of the experimental designs have been

used to conduct the experiments and the resulted data have been analyzed using standard statistical methods; however, little attention has been paid to the modeling of the effect of independent variables on the dependent ones. Recently, using efficient methods of the machine learning, model creation can be carried out easily. One of the branches of machine learning is classification. In classification, we want the model to look at features and then predict which category (formally called class), among some discrete set of options, an example belongs [9]. To conduct a machine learning modeling, the data that is given to the computer must have a special format called a data frame. In a data frame, each row is devoted to one data example, sometimes called data sample, and each column is

* Corresponding Author E-mail: i_ahmadi_m@yahoo.com

devoted to one property of the sample which is called feature. The last column of a data frame is a special feature of the sample which is called target variable. In machine learning, the target variable is abbreviated to y , and the other features are abbreviated to X . Furthermore, data features can be either continuous variables or categorical variables. If the data features are categorical, the problem is classification, otherwise the problem is regression. To perform classification, data samples must be stochastically divided into two groups named train dataset and test dataset. The data samples of the train dataset will be used to train the model, while the data samples of the test dataset will be used to test the model. On the other hand, all of the features of a classification problem must be converted to a special format called one-hot encoding or one-out-of- N encoding, also known as dummy variables. All of these divisions and conversions as well as modeling itself, can be done in the Python software which is a well-known platform to perform machine learning tasks.

In this study due to the scarceness of machine learning modeling in the field of soil science, a classification modeling was performed to classify different responses of *Zea Mays L.* as affected by different tillage operations performed and various soil amendments given to the soil.

2. Materials and Methods

2.1. Data

The raw material required for developing a machine learning model is data. Necessary data of this research was obtained from an open access source presented by Kiboi et al. ([10]). This source was composed of some Excel files containing the responses of *Zea Mays L.* (i.e. corn and stover yields, plant height at 6th and 10th leaf phases, and relative chlorophyll content of the crop leaves at 6th and 10th leaf phases) as affected by different tillage practices performed (i.e. conventional and minimum tillage), and different soil amendments used (i.e. C¹, F², RF³, RFM⁴, RTiP⁵, RML⁶, RTiM⁷). Research

experiments were conducted for two consecutive years (i.e. 2016 and 2017) in Chuka and Kandara sites in Kenya.

2.2. Method

In this study, all of the data features were categorical variables; therefore, the model considered herein is a classification model, and the number of classes can be adjusted by the number of labels assigned to the target variable. In other words, if two class classification is considered, values of the target variable should be divided into two classes. It can be performed using the median of the target variable values as the divider. The values above the median can be labeled as 1, and the values below the median can be labeled as 0. Or, if the four class classification is considered, the three Quartiles of the target variable values (i.e. Q1, Q2, and Q3) can play the divider roles (i.e. the values below Q1 can be labeled as 0, The values between Q1 and Q2 can be labeled as 1, the values between Q2 and Q3 can be labeled as 2, and the values above Q3 can be labeled as 3).

From the data features point of view, it should be noted that to analyze categorical variables, their values (which are strings instead of numbers) should be converted to one of the special formats considered for this purpose. The most common way to represent them is by using the one-hot encoding method. A one-hot encoding is a vector with as many components as we have categories in each variable. The component that corresponds to a particular instance's category is set to 1 and all other components are set to 0.

In order to examine the effect of different algorithms considered in the models (i.e. Decision Tree Classifier, Support Vector Machine (SVM) Classifier, K-Nearest Neighbors (KNN) Classifier, and Naive Bayes Classifier as standard Machine Learning (ML) algorithms, and Multi-Layer Perceptron (MLP) Classifier as a Deep Learning (DL) algorithm) on the model performance, classification accuracy and confusion matrix, as well as precision, recall and F1 score indicators were used as the model evaluation metrics. Accuracy is defined as the number of true predictions divided by total number of samples, while confusion matrix provides how many samples for each class are correctly classified and how many are confused with other classes. Precision is defined as the quality of a positive prediction made by the model. It refers to

¹ - Control

² - Sole mineral fertilizer

³ - Crop residue + mineral fertilizer

⁴ - Crop residue + mineral fertilizer + animal manure

⁵ - Crop residue + *Tithonia diversifolia L.* + phosphate rock (Minjingu)

⁶ - Crop residue + animal manure + legume intercrop (*Dolichos Lablab L.*)

⁷ - Crop residue + *Tithonia diversifolia L.* + animal manure

the number of true positives divided by the total number of positive predictions. Recall measures the model's ability to detect positive samples. It is calculated as the ratio between the number of positive samples correctly classified as positive to the total number of positive samples; therefore, the higher the recall indicator, the more positive samples detected. F1 score is calculated as the harmonic mean of precision and recall indicators, therefore, the F1 score integrates the precision and recall indicators into a single indicator in order to gain a better understanding of the model performance. If each of the comprising indicators of the F1 score diminishes, the nature of the harmonic mean makes the F1 score to reduce too; therefore, F1 score will be high if both precision and recall indicators are high.

Based on these criteria, the model having the best performance was selected as the final model.

2.3. Software

In this study, modeling was carried out in the Python software environment (Python Version: 3.7.6). The codes written for the corn yield modeling has been given in Appendix A. Similar codes were written for modeling of other target variables. The headlines of procedures followed in codes is summarized here:

- Importing necessary Python libraries
- Loading the data

- Labeling the values of the target variable of the Excel file based on the number of classes required for modeling
- Importing the modified Excel file into Python
- Performing preprocess operations on the data if necessary
- Converting data features into dummy variables
- Dividing the values of dummy and target variables into train and test datasets
- Introducing the model
- Training the model using the train dataset
- Calculating the classification accuracy and confusion matrix of the test dataset
- Printing the results

3. Results and Discussion

3.1 Effect of the Number of Classes on the Evaluation Metrics of Corn Yield Model

When SVM algorithm was used to classify corn yield values into classes and the number of classes increased from 2 to 4, the accuracy of the model reduced from 0.97 to 0.82. Figure 1 shows the confusion matrices of corn yield two-class (a), and four-class (b) classifiers.

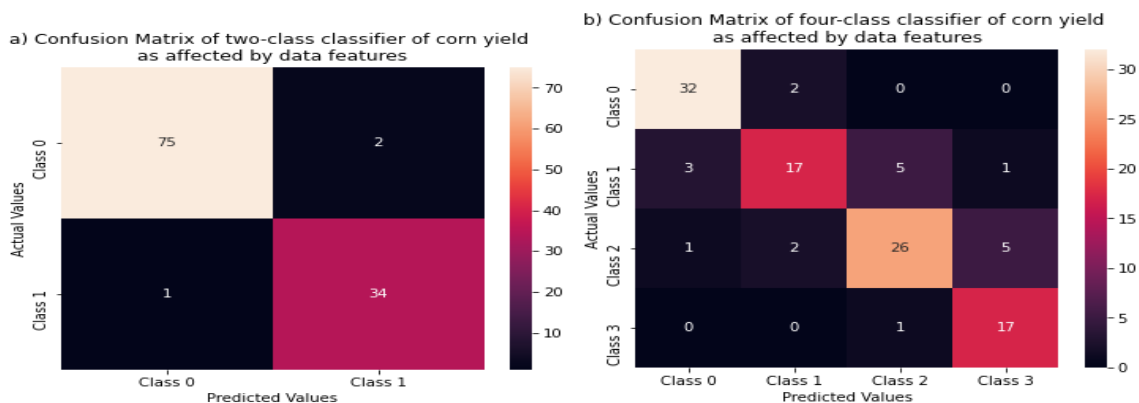


Fig 1. The confusion matrices of two-class (a) and four-class (b) classifiers of corn yield as affected by some independent variables

A good classifier is the one that predicts the class of each sample of the test dataset correctly; in other words, the higher the concentration of large numbers along the main diameter of the confusion matrix, the better. Moreover, as stated in the Introduction,

model evaluation is performed on the samples of the test dataset; therefore, cell values of both matrices sums cumulatively to 112, because the test dataset considered herein contains 112 samples (Note that the total number of samples was 448, thus the

Python has separated them stochastically to 336 samples of the train dataset and 112 samples of the test dataset). Another clear point is that the increase in the number of classes is achieved by sacrificing 12% of the model accuracy. Therefore, there is always a trade-off between the number of classes and accuracy of the model, and the researcher must take this compromise into account.

3.2. Modeling Accuracies of Two-Class and Four-Class Classifications of Target Variables as Affected by Different Classification Algorithms

Table 1 shows the values of two-class and four-class accuracies obtained from applying four different algorithms (i.e. Decision Tree Classifier, SVM Classifier, KNN Classifier, and Naive Bayes Classifier) for classifying the target variables considered in this research (i.e. crop yield, stover yield, crop height at 6th and 10th leaf phases, and relative chlorophyll content of the crop leaves at 6th and 10th leaf phases).

Table 1.

Modeling accuracies obtained from applying four different algorithms for classifying the target variables considered in this research

Algorithm name	Modeling accuracy of corn yield		Modeling accuracy of stover yield		Modeling accuracy of crop height at 6th leaf phase		Modeling accuracy of crop height at 10th leaf phase		Modeling accuracy of relative chlorophyll content at 6th leaf phase		Modeling accuracy of relative chlorophyll content at 10th leaf phase	
	Two class	Four class	Two class	Four class	Two class	Four class	Two class	Four class	Two class	Four class	Two class	Four class
Decision Tree	0.76	0.56	0.64	0.41	0.66	0.42	0.88	0.63	0.76	0.5	0.69	0.4
SVM Classifier	0.97	0.82	0.83	0.5	0.7	0.54	0.92	0.75	0.77	0.54	0.79	0.56
KNN Classifier	0.88	0.63	0.78	0.5	0.69	0.49	0.91	0.73	0.78	0.47	0.7	0.44
Naive Bayes	0.76	0.55	0.74	0.41	0.76	0.45	0.85	0.54	0.77	0.47	0.77	0.39

As shown, the application of the SVM classifying algorithm led to relatively higher modeling accuracies; therefore, the SVM algorithm was selected as the most appropriate algorithm in this

research. Furthermore, the four-class confusion matrices of SVM modeling of the other target variables (the confusion matrix of crop yield has been shown in figure 1) are depicted in figure 2.

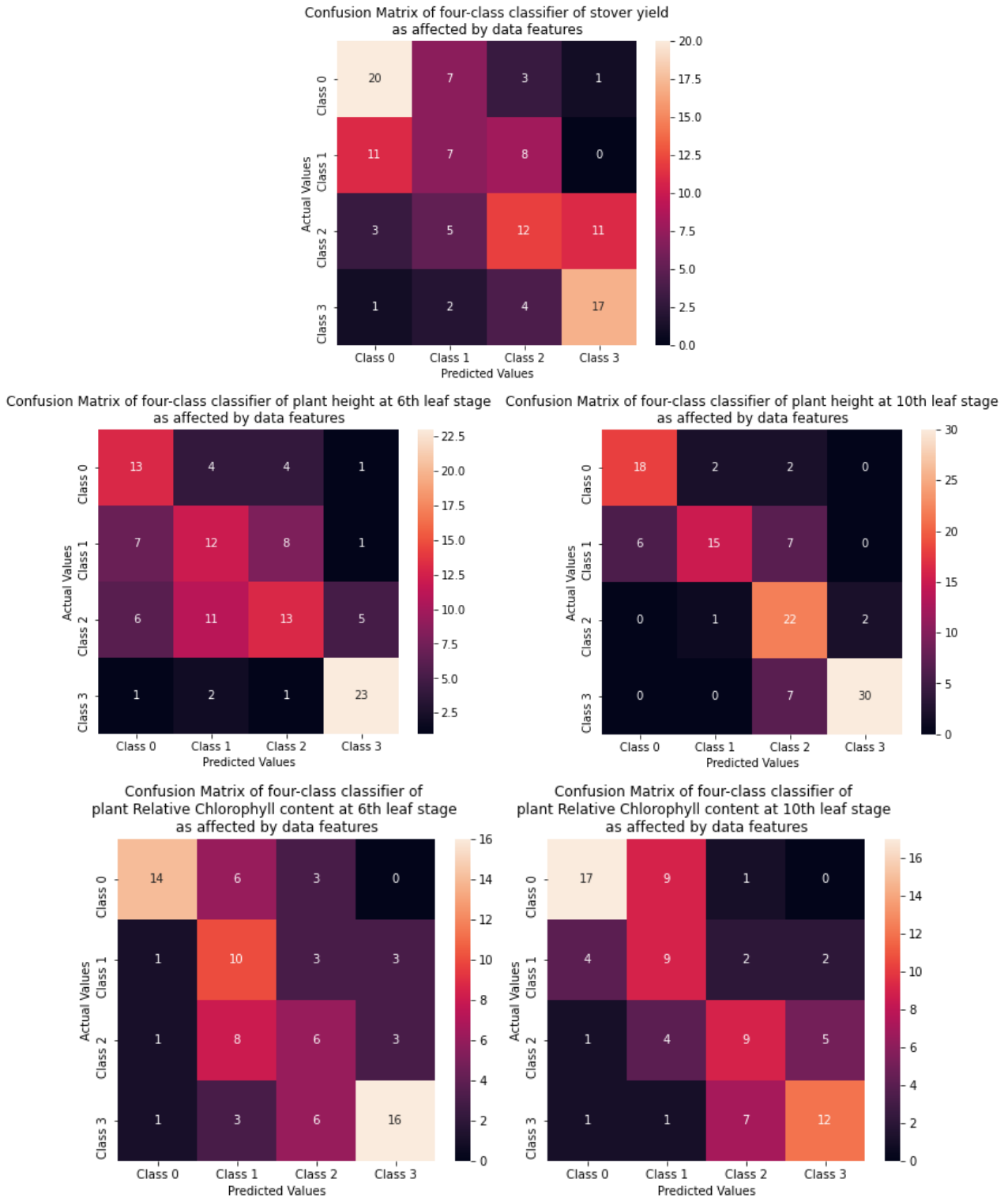


Fig 2. Confusion matrices of SVM classification of stover yield, crop height at 6th and 10th leaf phases, and relative chlorophyll content of the crop leaves at 6th and 10th leaf phases

In the case of relative chlorophyll content of the crop leaves at 6th and 10th leaf phases, the values within matrices sums cumulatively to 84 instead of 112, because the raw data for these two target variables

were composed of 336 samples; therefore, one fourth of this number that comprises the number of the test dataset will be 84.

3.3. Indicator Values of Two-Class and Four-Class Classifications of Target Variables as Affected by the MLP Algorithm

Tables 2 and 3 show the indicator values obtained for two-class and four-class classification of dependent variables that were affected by the MLP algorithm:

Table 2.

Performance indicators obtained for two-class classification of dependent variables using the MLP algorithm

	class	precision	recall	F1 score	support	accuracy
Corn yield	0	0.96	1	0.98	43	0.98
	1	1	0.96	0.98	47	
Plant height at 10th leaf phase	0	0.94	1	0.97	45	0.97
	1	1	0.93	0.97	45	
Plant height at 6th leaf phase	0	0.86	0.96	0.91	51	0.89
	1	0.94	0.79	0.86	39	
Relative chlorophyll content at 10th leaf phase	0	0.83	0.78	0.81	32	0.82
	1	0.82	0.86	0.84	36	
Relative chlorophyll content at 6th leaf phase	0	0.87	0.84	0.85	31	0.87
	1	0.87	0.89	0.88	37	
Stover yield	0	0.85	1	0.92	46	0.91
	1	1	0.82	0.9	44	

Table 3.

Performance indicators obtained for four-class classification of dependent variables using the MLP algorithm

	class	precision	recall	F1 score	support	accuracy
Corn yield	0	1	0.79	0.88	19	0.92
	1	0.83	1	0.91	24	
	2	1	0.87	0.93	23	
	3	0.92	1	0.96	24	
Plant height at 10th leaf phase	0	0.73	1	0.84	19	0.87
	1	0.96	0.88	0.92	26	
	2	1	0.62	0.77	24	
	3	0.84	1	0.91	21	
Plant height at 6th leaf phase	0	0.71	0.94	0.81	32	0.76
	1	0.77	0.42	0.54	24	
	2	0.79	0.86	0.83	22	
	3	0.82	0.75	0.78	12	
Relative chlorophyll content at 10th leaf phase	0	0.75	1	0.86	18	0.79
	1	1	0.43	0.61	23	
	2	0.72	0.93	0.81	14	
	3	0.81	1	0.9	13	
Relative chlorophyll content at 6th leaf phase	0	0.71	0.88	0.79	17	0.74
	1	0.65	0.93	0.76	14	
	2	1	0.45	0.62	22	
	3	0.71	0.8	0.75	15	
Stover yield	0	0.96	0.92	0.94	26	0.92
	1	1	0.95	0.97	20	
	2	0.91	0.84	0.87	25	
	3	0.83	1	0.9	19	

Comparison of the results shown in tables 2, and 3 with the results presented in table 1, shows that the deep learning method considered herein i.e. the MLP

has higher performance indicators compared to the indicators obtained for the standard machine learning methods.

3.4. Procedure for using the Model

After the establishment of a machine learning model, the relationship between the model inputs and output is known for the computer; therefore, if the model inputs are given to the computer as a vector, computer can convert it to the output value, using the model established. In a classification problem, the input vector is composed of a string of zeros and ones, which its components can be categorized according to the dummy variables used to establish the model. In the present study, because 5 dummy variables with 19 binary locations were used, there are five 1s in the vector (exactly one 1 for each dummy variable), and the other 14 binary locations filled with 0s. For example in the string [1,0, 1,0,0,0, 1,0, 1,0,0,0,0,0,0, 1,0,0,0], the first 2 binary locations belongs to the dummy variable "Experiment Site" (Chuka and Kandara), the next 4 binary locations belongs to the dummy variable "Experiment Season" (Season_LR 2016, Season_SR 2016, Season_LR 2017, and Season_SR 2017), the next 2 binary location belongs to the dummy variable "Tillage" (Conventional and Minimum tillage), the next 7 binary locations belongs to the dummy variable "Soil amendments" (C, F, RF, RFM, RTiP, RML, and RTiM), and the last 4 locations belongs to the dummy variable "Rep" (Rep_1, Rep_2, Rep_3, and Rep_4). Using the Python codes: `A = svm_model_linear.predict([[1,0,1,0,0,0,0,1,1,0,0,0,0,`

`0,0,1,0,0,0]]), and print(A); one can give a vector like above to the software and obtain the model output accordingly. Table 4 shows the classes obtained from the modeling of corn yield as affected by tillage type and soil amendments:`

Table 4
Classes obtained from the modeling of corn yield as affected by tillage type and soil amendments

Soil amendment	C	F	RF	RFM	RML	RTiM	RTiP
Conventional tillage	2	3	3	3	2	3	3
Minimum tillage	2	3	3	3	2	3	3

Table 4 shows that the type of tillage has not any effect on the crop yield, while soil amendments C and RML resulted in a lower yield than the other soil amendments. These results are in agreement with the results obtained by Kiboi et al. ([11]), who reported from ANOVA that tillage type had not any significant effect on crop yield, while soil amendments had a significant effect on this target variable, and the comparison of means showed that the C and RML treatments resulted in a lower yield in comparison with the other treatments.

A more complete representation of table 4 that separates the role of experiment sites and cultivation seasons on the value of corn yield, are shown in table 5 and 6. Table 5 is for the Chuka site, and table 6 is for the Kandra site.

Table 5
Detailed representation of the scattering of corn yield classes over the independent variables' matrix in the Chuka experimental site

Cultivation season	Type of tillage	Soil amendments						
		C	F	RF	RFM	RML	RTiM	RTiP
LR 2016	Conventional	2	3	2	3	2	3	3
	Minimum	2	3	2	3	2	3	3
LR 2017	Conventional	3	3	3	3	2	3	3
	Minimum	2	3	3	3	2	3	3
SR 2016	Conventional	1	2	2	2	1	2	1
	Minimum	1	2	2	2	0	2	1
SR 2017	Conventional	0	1	1	1	0	1	0
	Minimum	0	1	1	1	0	1	1

Table 6

Detailed representation of the scattering of corn yield classes over the independent variables' matrix in the Kandra experimental site

Cultivation season	Type of tillage	Soil amendments						
		C	F	RF	RFM	RML	RTiM	RTiP
LR 2016	Conventional	2	3	2	3	2	3	3
	Minimum	2	3	3	3	2	3	3
LR 2017	Conventional	0	0	0	0	0	0	0
	Minimum	0	0	0	0	0	0	0
SR 2016	Conventional	1	2	2	2	1	2	1
	Minimum	1	2	2	2	0	2	1
SR 2017	Conventional	2	2	2	2	1	1	1
	Minimum	1	2	2	2	1	1	1

Some results that can be inferred from table 5 and 6 are as follows:

- Both ANOVA and machine learning analysis showed that the crop yield values in the SR seasons had been lower than the crop yield values in the LR seasons (Hint: 0s obtained for the row devoted to the cultivation season LR of the year 2017 in the Kandra site are due to the lack of data in that season).
- Both ANOVA and machine learning analysis indicated higher yields obtained from mineral fertilization for cultivation season SR 2017 in the Kandra site compared to the other soil treatments.
- The ANOVA and machine learning analysis showed complete similarity regarding the higher yields obtained from mineral fertilization and RTiM treatment for cultivation season SR 2016 in the Kandra site compared to the other soil treatments.
- Both ANOVA and machine learning analysis predicted lower corn yields for RLM and Control treatments for cultivation season LR 2016 in the Kandra site compared to the other soil treatments.
- The ANOVA, similar to machine learning analysis, showed that the least value of corn yield in the Chuka experimental site has been obtained in the cultivation season SR 2017.
- The ANOVA and machine learning analysis showed complete similarity regarding the lower yields obtained from Control, RML, and RTiP treatments for cultivation season SR 2016 in the Chuka site compared to the other soil treatments.

These results act as cross checking for the appropriateness of the model developed in this study.

4. Conclusion

In this study the possibility of classifying different attributes of the response of Zea Mays L. to the combination of performing some tillage practices and applying some soil amendments using the machine learning principles was examined. It was shown that among the algorithms used, the Support Vector Machine (SVM) led to the best performance; however, its accuracy decreased as the number of classes increased. Therefore, there is a trade-off between classification accuracy and the number of classes allocated for Zea Mays L. modeling in this study. The procedure for using the model developed herein was also practiced in this paper. Finally, application of the MLP algorithm to classify each of the dependent variables considered herein, resulted in higher accuracies compared to the accuracies of the other standard ML algorithms.

5. Author Contribution

The raw data of this research was obtained from an open access source presented by Kiboi et al. ([10]). Data screening and modifying, modeling, discussing the results and drafting the manuscript were done by the single author of this manuscript.

6. Appendix A

Python codes for classifying corn yield values as a function of experiment location, experiment season, type of tillage practice, and type of soil amendments:


```
# importing necessary libraries
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.metrics import confusion_matrix
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.svm import SVC
from sklearn.neighbors import KNeighborsClassifier
from sklearn.naive_bayes import GaussianNB

#Loading the data
#Importing the modified Excel file into Python
data = pd.read_csv("Directory and name+suffix of the excel file
must be inserted here",header=0, index_col=False,

names=['Site','Season','Tillage','Soil_Input','Rep','Yield_Mg/ha'])
data =
data[['Site','Season','Tillage','Soil_Input','Rep','Yield_Mg/ha']]

#Converting data features into dummy variables
data_dummies = pd.get_dummies(data)

#Dividing the values of dummy and target variables into train
and test datasets
features = data_dummies.loc[:, 'Site_Chuka':'Rep_two']
X = features.values
y = data_dummies['Yield_Mg/ha'].values
X_train, X_test, y_train, y_test = train_test_split(X, y,
random_state = 0)

#Introducing the model
dtree_model = DecisionTreeClassifier(max_depth =
2).fit(X_train, y_train)

#Training the model using the train dataset
dtree_predictions = dtree_model.predict(X_test)

#Calculating the classification accuracy and confusion matrix of
the test dataset
accuracy = dtree_model.score(X_test, y_test)
cm = confusion_matrix(y_test, dtree_predictions)

#Printing the results
print(accuracy)
cm_df = pd.DataFrame(cm,
index = ['1st','2nd','3rd','4th'],
columns = ['1st','2nd','3rd','4th'])
plt.figure(figsize=(6,5))
sns.heatmap(cm_df, annot=True)
plt.title('Confusion Matrix')
plt.ylabel('Actual Values')
plt.xlabel('Predicted Values')
plt.show()

#Required code for using the model
A =
svm_model_linear.predict([[1,0,1,0,0,0,0,1,1,0,0,0,0,0,0,1,0,0]
])
print(A)
```

References

- [1] Tesfahunegn, G. B. 2015. Short-term effects of tillage practices on soil properties under Tef [Eragrostis tef (Zucc. Trotter)] crop in Northern Ethiopia. *Agric. Water Manag.* 148: 241–249. <https://doi.org/10.1016/j.agwat.2014.10.004>.
- [2] Kiboi, M. N., K. F. Ngetich, J. Diels, M. Mucheru-Muna, J. Mugwe and D. N. Mugendi. 2017. Minimum tillage, tied ridging and mulching for better maize yield and yield stability in the Central Highlands of Kenya. *Soil Tillage Res.* 170: 157–166. <https://doi.org/10.1016/j.still.2017.04.001>.
- [3] Lazcano, C., M. Gómez-Brandón, P. Revilla and J. Domínguez. 2013. Short-term effects of organic and inorganic fertilizers on soil microbial community structure and function. *Biol. Fertil. Soils.* 49: 723–733. <https://doi.org/10.1007/s00374-012-0761-7>.
- [4] Ibrahim, A., R. Clement, D. Fatondji and A. Opoku. 2015. Hill placement of manure and fertilizer micro-dosing improves yield and water use efficiency in the Sahelian low input millet-based cropping system. *F. Crop. Res.* 180: 29–36. <https://doi.org/10.1016/j.fcr.2015.04.022>.
- [5] Chivenge, P., B. Vanlauwe, R. Gentile and J. Six. 2011. Comparison of organic versus mineral resource effects on short-term aggregate carbon and nitrogen dynamics in a sandy soil versus a fine textured soil. *Agric. Ecosyst. Environ.* 140: 361–371. <https://doi.org/10.1016/j.agee.2010.12.004>.
- [6] Mucheru-Muna, M., D. Mugendi, P. Pypers, J. Mugwe, J. Kung'u, B. Vanlauwe and R. Merckx. 2014. Enhancing maize productivity and profitability using organic inputs and mineral fertilizer in Central Kenya small-hold farms. *Exp. Agric.* 50: 250–269. <https://doi.org/10.1017/S0014479713000525>.
- [7] Paul, B. K., B. Vanlauwe, F. Ayuke, A. Gassner, M. Hoogmoed, T. T. Hurisso, S. Koalab, D. Leleib, T. Ndabamenyea, J. Six, M. M. Pulleman. 2013. Medium-term impact of tillage and residue management on soil aggregate stability, soil carbon and crop productivity. *Agric. Ecosyst. Environ.* 164: 14–22. <https://doi.org/10.1016/j.agee.2012.10.003>.
- [8] Grabowski, P. P., S. Haggblade, S. Kabwe and G. Tembo. 2014. Minimum tillage adoption among commercial smallholder cotton farmers in Zambia, 2002 to 2011. *Agric. Syst.* 131: 34–44. <https://doi.org/10.1016/j.agsy.2014.08.001>.
- [9] Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. 2021. Dive into deep learning. Release 0.16.1.
- [10] Kiboi, M., A. Fliessbach, A. Muriuki and F. Ngetich. 2022. Data on the response of Zea Mays L. and soil moisture content to tillage and soil amendments in the sub-humid tropics. *Data Br.* 43: 108381. <https://doi.org/10.1016/j.dib.2022.108381>

- [11] Kiboi, M. N., K. F. Ngetich, A. Fliessbach, A. Muriuki and D. N. Mugendi. 2019. Soil fertility inputs and tillage influence on maize crop performance and soil water content in the Central Highlands of Kenya, *Agric. Water Manag.* 217: 316–331. <https://doi: 10.1016/j.agwat.2019.03.014> .