Mapana - Journal of Sciences
2023, Vol. 22, No. 3, *81-98*
ISSN 0975-3303 | https://doi.org/10.12723/mjs.66.4

An Ensemble Learning Methodology to a Decision Tree Algorithm for Soil Type Classification Using Machine-learning

Sumanth. S.*

Abstract

In this study, we focus on the classification of soil types in a specific region, employing a stacking ensemble learning approach with the decision tree algorithm. The importance of this research stems from the fact that a significant portion of the population in rural areas relies heavily on agriculture for their livelihood, making agriculture the backbone of our nation's economy. In India, a country with diverse soil types, understanding soil characteristics is critical for agricultural development as it profoundly influences crop productivity. Data mining techniques play a pivotal role in predicting soil types assisting farmers in selecting the most suitable crops for cultivation. To address this agricultural challenge, we propose a novel method named "Stacking Ensemble Learning with the Decision Tree Model" for soil type classification. Our approach outperforms existing Decision Tree-based methods and exhibits unique advantages in solving complex soil classification problems. Our experimental results demonstrate the effectiveness of our approach in creating an optimal decision tree model for soil type classification. Specifically, our study reveals that the Ensemble Learning technique, leveraging stacking with the Decision Tree model, offers more accurate soil type classification compared to several machine-learning algorithms employed in this research, such as K-Nearest Neighbours (KNN), Support Vector Machine (SVM), traditional

* Department of Computer Science and Applications, Government College for Women, Kolar, Karnataka, India – 563 101; Email: sumanth81s@gmail.com

Decision Tree, and a Bayesian approach to the Decision Tree model.

Keywords: Soil Types, Crop Productivity, Decision Tree, Ensemble Learning, Machine-learning

1. Introduction

A big dataset is being analysed using machine-learning techniques to create classification models and find valuable trends. The foundation of machine-learning is computer science and statistics, which allows computers to develop models from sample data and automate decision-making based on the provided soil dataset. Machine-learning employs a model to identify patterns in incoming data, predict outcomes, and learn from existing data patterns. Understanding data organisation and incorporating it into models that people can use and understand is the aim of machine-learning. Artificial intelligence is the term used to describe the replication of human intelligence in devices that have been programmed to think and behave just like people. Datasets are subjected to data preprocessing procedures to get them ready for the efficient classification process."

"Large dataset categorization often uses the influential decision tree technique. It is a member of the family of supervised machine-learning. It makes it simple for us to comprehend and interpret datasets. Building a training model that could be used to predict the class or values of the target variable is the aim of employing a decision tree. Each branch of a decision tree in the decision tree technique corresponds to an attribute, and each leaf node to the class label. The efficacy and correctness of the soil datasets gathered from a specific area are evaluated. In order to demonstrate that our suggested algorithm provides a more valuable and accurate result than the existing models, the results of existing algorithms such as KNN, SVM, Decision Tree, and Bayesian method are used in the decision tree model."

Soil can be classified from the viewpoints of a resource and a material. Commercial laboratories that provide a variety of tests, each focusing on a certain set of soil minerals and chemicals, can perform soil testing. The advantage of using a soil testing lab is that they are familiar with the chemistry of the soil in the area from which the samples are gathered. This makes it possible for specialists to suggest the tests that are most likely to produce valuable data on the soil type. In our country, there are eight main classifications of soil types. "They are Alluvial soils, Black soils, Red soils, Laterite and Lateritic soils, Forest and Mountain soils, Arid and Desert soils, Saline and Alkaline soils, and Peaty and Marshy soils. However there are 8 major categories of soil, only 6 categories are considered in our research work, since there can be a maximum of 6 types of soil present in a particular region.

In this work, K-Nearest Neighbour, Support Vector Machine, Decision Tree, and Bayesian method for a decision tree model are some of the pre-existing approaches that we use a machine-learning tool to effectively and accurately reveal the findings in. The proposed method's new algorithm has been developed, and it is implemented by writing Python source code. The results of the existing and suggested methods are shown in Table 2. With the use of bar charts, the results are both recommended and compared to current algorithms.

2. Related Works

Alobaidi et al. [1], who explored a few practical ensemble learning models, a thorough examination of eight ensemble models may be carried out utilising five machine-learning algorithms and three ensemble techniques. As compared to a single model, the outcome was found to improve performance prediction and diminish the uncertainty of ensembles.

Ramesh Babu Palepu et al. [2], investigated agricultural soil for resolution-making in numerous issues related to the agricultural domain. In order to analyse the type of soil and its characteristics, which will be more acceptable to develop the right crop, several data mining tools and methodologies were presented. The accuracy of the results increases as the dataset size increases, and this may also make it easier to identify appropriate patterns when compared to traditional statistical analysis.

P.Surya et al. [3], tested various regression techniques with the help of collected agriculture soil datasets and devised an algorithm for the effective prediction of agricultural yield for various crops in Tamilnadu, particularly in and around Erode district, based on soil type classification. Additionally, discuss several algorithms, such as ID3, C4.5, and CART, and contrast their performance and results. The available datasets were used to help in their evaluation. Different decision tree algorithms are evaluated for performance based on their confidence and the amount of time needed to see the tree. Out of the three algorithms discussed in this study, it has been shown that C4.5 offers the required precision and proficiency in comparison to other algorithms, making it the best algorithm.

"To address the shortcomings in the current decision tree models, such as the C4.5 and CART algorithms, Vrushal Milan Dolas et al. [4], suggested an improved Decision Tree method. When the domain of the target attribute is relatively big, the CART method creates misclassification errors and the C4.5 is biased towards characteristics with more values. The suggested approach aids in precisely classifying the soil type from the big dataset of a specific location that has been provided."

"A technique that incorporates verifying the soil quality to forecast the right crop to be grown according to the soil type was provided by Rushika Gadge et al. [5]. According to the kind of soil, the farmer cultivates crops that are appropriate for the soil and uses the prescribed fertilisers to optimise crop yield, which is greatly helped by the algorithm provided in this study."

"In order to forecast the soil types from the provided soil dataset of a specific location, N. Saranya et al. [6], experimented with machine-learning classification techniques such as K-Nearest Neighbour, Bagged tree, Support vector machine, and logistic regression. After using algorithms to extract information from soil data, two types of soil, such as Black and Red soil, are taken into account for optimal crop cultivation. Agriculturalists can grow the right crop depending on the anticipated soil type by using the aforementioned classification algorithms."

"Chandan et al. [7], investigated methods for classifying different types of soil based on their moisture, nutrients, structure, quality, pH, and texture using decision trees, support vectors, and K-Nearest Neighbour algorithms from the provided soil dataset."

"In an effort to classify soil types using data mining and pattern recognition techniques, Dr. D. Ashok Kumar et al. [8], conducted a survey. They came to the realisation that efficient categorization techniques were needed in order to assist agriculturalists in cultivating crops that would be more profitable for them."

"In their study of the significance of data mining techniques, Ashwini Rao et al. [9], discovered that a variety of algorithms and strategies are utilised to categorise soil for the growing of suitable crops. Also, they discovered how to improve the current Support Vector Machine-learning technique, which they used to classify a sizable soil dataset. The suggested algorithms take into account and process the coloured image of the soil sample in order to extract various characteristics, including colour, texture, and different soil types, such as red, black, clay, and alluvial."

P. A. Harlianto et al. [10], The categorization of soil types may be automated using a machine-learning system. This study analyses several machine-learning classification techniques for soil types. For this classification, support vector machine (SVM), neural network, decision tree, and naive Bayesian algorithms are suggested and evaluated. The soil dataset is derived from actual data. RapidMiner Studio is used to execute the simulation. The correctness is shown by the performance. The result demonstrates that SVM outperforms other methods when linear function kernels are used. 82.35% is the SVM's best accuracy.

N. Ajithkumar et al. [11], The categorization of landmines and clutter is one component of the multiphase robot-based landmine detection challenge. The inability to identify a landmine, the length of the detection process, and the significant number of false alarms that result from faulty classification must all be taken into account while designing an efficient and successful classification model. This study intends to analyse 5 distinct classifiers, namely: Hidden Markov Model, Support Vector Machine, Artificial Neural Network, Gradient Boosted Decision Tree, and Adaptive Boosted Decision trees, in the absence of a comprehensive analysis of the efficacy of such models. Two open-source datasets based on GPR that each offer information on arid, desert-like soils and vegetation have been utilised. Different ratios of the mine to non-mine data are taken into consideration in order to make the research complete in terms of class label percentage as well. Confusion matrices and their accompanying measurements have been used to compare the models. Based on this, a selection table has been created that enables the user to choose the classifier that is most likely to perform well in relation to a selected measure and the training dataset that they would like to use.

S. Raskar et al. [12], Predicting crop output is a very complex operation that depends on a wide range of variables, including state, district, region, rainfall, temperature, soil characteristics, and many more. We must determine the functional connection between the yield and the aforementioned elements in order to anticipate yields with accuracy. To find such links, large datasets and powerful algorithms are needed. This research primarily focuses on using different machine-learning approaches to forecast agricultural productivity and provide crop recommendations. Here, models like Decision Trees and Random Forests will be employed. By taking into account variables like temperature, rainfall, area, and other aspects, the predictions provided by these models will assist farmers in choosing which crop to cultivate to induce the greatest yield. A crop recommendation system that takes into account variables like rainfall, annual temperature, soil content, and type, among others, is crucial because farmers frequently grow the same crop on the same soil for years at a time, depleting the soil's nutrients in the process. This system also helps farmers earn a profit.

Y. Garg et al. [13], The damage increases with the earthquake's magnitude. Due to most bridges being constructed on soft soil, there is a greater likelihood that they may behave like ships in the sea during an earthquake due to the increased mobility of the soft soil. The most important responsibility to prevent any tragedies is to ensure the stability of the bridges. Because its mobility and sustainability cannot withstand an earthquake of this size, many bridges fall during one. In this article, we present a method to predict whether a bridge will be damaged or not after an earthquake by taking into account variables such as the magnitude of the earthquake, the distance from the bridge's epicentre, the type of bridge, the material used to construct the bridge, and many others. We used a variety of classification algorithms, including Logistic Regression, Decision Tree, Random Forest, XGBoost, and KNN. This forecast will assist in increasing the resilience of bridges during an earthquake, which will help to save many lives.

S. Bhansali et al. [14], India's economy is heavily dependent on agriculture. Agriculture still uses conventional methods of advice. Farmers now approximate the quantity of fertiliser used and the kind of crop to be seeded using conventional methods. The region's climate and soil type have a significant impact on agriculture. Therefore, it is crucial to foster progress in this area. We will develop a web application using machine-learning and deep learning techniques that will serve as a one-stop shop for agricultural information. The farmers would be assisted in raising their yield generation by the crop and fertiliser advice system. To choose the best crop for that area, we will consider the soil characteristics and the weather API. We will create a recommendation model using the decision tree and the navies bayes algorithm, using the N-K-P, Ph. value, and rainfall as the training parameters. We will suggest fertiliser and applications based on the crop and farming area to increase farmers' yield productivity. Crops may get sick at times as a result of unwelcome surplus rain or insect attacks. The system will identify the kind of illness using Support Vector Machine (SVM) or neural network approaches utilising the image classification methodology, which allows the user to input a photo of the damaged plant or crop. This disease identification will provide recommendations for how to treat or avoid that plant or crop. The goal is to create a single system for all the attributes and offer the most accurate results possible for all the crops in the majority of India's regions. Additionally, the pricing and news area will keep farmers informed about current market rates as well as government programmes and policies pertaining to agriculture and farming.

R. Kumar et al. [15], The standard quality and quantity of agricultural output are severely harmed by the widespread occurrence of crop diseases and by the inability of the soil to support plant growth. Therefore, it is important to identify crop diseases as soon as possible by coming up with or using a quick, creative method, and a crop recommendation system will be helpful to farmers. As a result, this research suggested a system that uses CNN to identify plant illnesses and uses ML to analyse the soil's numerous attributes to suggest different crops depending on the soil's quality. The dataset used for disease prediction training and testing was successfully split from the Plant Village Dataset, allowing different plant species to be identified and given new names to create an accurate database. The next phase is to acquire a test database made up of several plant diseases that will be used to evaluate the proposed module's precision and degree of confidence. The classifier is then trained using training data, and the output will then be identified with the highest accuracy possible. To increase the effectiveness of our model, we employ the Support Vector Classifier (SVC) method for the crop recommendation system, which beats other classifiers like KNN, Logistic Regression, Random Forest, and Decision Trees. The created model also maps the database of soil and crops and recommends appropriate crops depending on the amount of nutrients present in the soil, enabling farmers to choose the right crops to plant in their fields. Additionally, this research assessed the effectiveness of multiple classifiers on the study's dataset and selected the one with the greatest accuracy.

3. Proposed Model

The suggested model employs a unique approach that outperforms existing algorithms in accurately classifying soil types. The core of this approach is the utilization of Stacking with Ensemble Learning integrated with the Decision Tree model. Python is the chosen programming language for the design and implementation of this algorithm. We evaluate our model's efficacy by contrasting its output with findings from various machine-learning methods, such as the Decision Tree model using a Bayesian approach, K-Nearest Neighbour, Support Vector Machine, and conventional Decision Tree.

The soil dataset provided for this study undergoes a systematic process of segmentation. Our novel method partitions the dataset into smaller and more specific subgroups. This segmentation process involves various sorting methods, resulting in a reduction in the number of samples in each partition. The principle underlying this process is recursive data partitioning, wherein the dataset is progressively divided into more homogeneous subsets. These subsets are referred to as nodes once the partitioning is complete, and each leaf node is assigned the label of the majority class within it.

At the heart of our approach lies the Decision Tree algorithm, a fundamental technique in machine-learning and predictive modeling. Decision trees employ a "divide and conquer" strategy to split data based on attribute values. When making predictions using a decision tree for a given record, the process starts at the tree's root node. The values of the root attributes are compared with the attributes of the record. Based on this comparison, the process follows the branch corresponding to the attribute value and moves to the next node, repeating this process until reaching a leaf node.

Ensemble Learning

A single ultimate predictive model is created via ensemble learning, a type of machine-learning that incorporates multiple base models. In many classification situations, the ensemble learning methodology could be used to get better results than any other strategy. The predictions of a selected group of hypotheses (an ensemble) are combined in ensemble learning. Several basic learners are created using the ensemble learning approach, and when they are joined, multiple learning models (classifiers) are solidly built for the classification of data objects for the supplied dataset.

The Ensemble Learning Method aims to enhance the performance of the prediction model by increasing the findings' accuracy and efficiency in the fields of statistics and machine-learning. Several models are combined using the ensemble learning technique, which increases predictive power. For instance, when group members come from different disciplines, the group is likely to make better decisions than the individual members.

“Ensemble learning's core idea is to get multiple-base learners ready for group participation and combine their predictions into a single output that, on average, should perform better. By mixing many base models, ensemble learning enhances the outcomes. In order to minimise variance (Bagging), bias (Boosting), and enhance prediction (Stacking), ensemble learning approaches integrate many Machine-learning techniques into a single predictive model. Stacking differs from bagging and boosting primarily in two ways. First, whereas boosting primarily considers homogeneous weak learners, stacking frequently takes into account heterogeneous weak learners (multiple learning methods are merged). Second, bagging and boosting combine weak learners, whereas stacking learns to combine the basic model using a meta-model.”

To increase the effectiveness and accuracy of the outcome in this study, stacking ensemble learning was applied to the decision tree model. A meta-classifier is used in the stacking technique of ensemble learning to merge different categories. The meta-models are trained using the output of the base-level model-like features after the base-level models have been trained using a large training set. Because there are several learning techniques at the base level, stacking ensembles are heterogeneous.

Proposed Method

In this study, the "soil.csv" soil dataset, which includes 5200 instances and the attributes (OC, P, Fe, Mn, Cu, K, pH, Zn, and EC), is utilised as input to both current and new algorithms. “In order to develop a classifier model for the prediction, 80% (4160 instances) of the training sample and 20% (1040 instances of the uncertain class label) of the test data are first used. The soil data’s properties are mentioned in Table-1.

Attribute or Fields	**Description**
OC	Organic Carbon, %
P	Phosphorous, ppm (parts per million)
Fe	Iron, ppm
Mn	Manganese, ppm
Cu	Copper, ppm
K	Potassium, ppm
pH	pH value of the soil
Zn	Zinc, ppm
EC	Electrical conductivity, decisiemen per meter

Table-1: Attribute descriptions of the Soil dataset

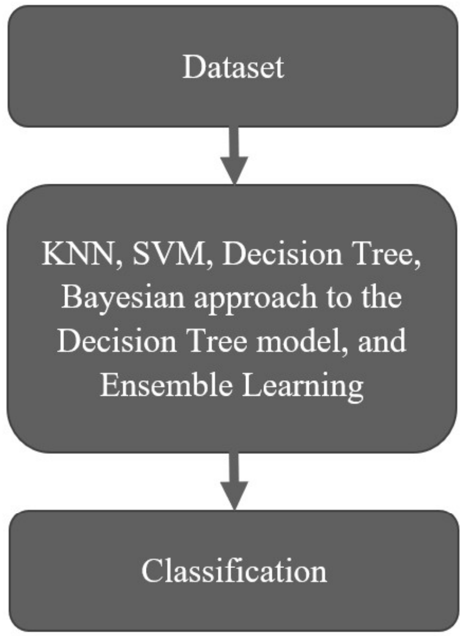

Figure-1: Soil classification using the Proposed Method (Ensemble learning method)

The soil dataset may be divided into a training dataset and a testing dataset using the Python source code below:

“import pandas as pds;

from sklearn.model_selection import train_test_set_split; Dataset1 = pds.read_csv (“soil.csv”);

Dataset1.head ( );

x_train_set, y_test_set, y_train_set, y_test_set=train_test_set_split (x, y, test_size=0.2);

x_train_set.head ( );

Steps for soil classification using Stacking Ensemble Learning with Decision Tree Model are represented below:

Input: Soil dataset DS

Output: Expected class of Soil types.

Step 1: The input is given as dataset DS and is considered as samples for soil type prediction Step 2: The given soil dataset DS has to be preprocessed for our research need.

Step 3: The feature extraction process will be performed, which will reduce the dimensionality of the data given as input. Then the given dataset is trained and tested.

Step 4: Ensemble Learning using the Stacking method is used with a Decision Tree.

a) Start the tree with the root node, which contains the complete dataset
DS

b) Find the best attribute in the dataset using Attribute Selection Measure (ASM)

c) Divide the given data set DS into subsets that contain possible values for the best attribute

d) Generate the decision tree node, which contain the best attribute.""

Step 5: "For Stacking K-Nearest Neighbor, Random Forest and XGB classifications are used.

Step 6: Recursively make new decision trees using the subsets of the dataset created in step 3 and 4 and continue this process until a stage is reached where we cannot further classify the nodes and the final node is a leaf node.

Finally, tree formation is done, soil type is predicated, the performance analysis is carried out using precision, recall, F1-Score, and accuracy.

4. Results and Discussion

With the use of a robust database of Python libraries, the suggested algorithm for classifying soil types is put into practice. The findings are reported in Table 2, which is displayed below. The four metrics employed in the proposed algorithm precision, recall, F1-score, and accuracy are connected to the categorization of soil type from a specific region's soil dataset. Last but not least, these four metrics provide the framework for evaluating the classification task. The two key measures for evaluating a classifier's output characteristics are precision and recall. Precision (also known as positive predictive value) refers to the proportion of relevant results and Recall (also known as sensitivity) refers to the percentage of total relevant results correctly classified by our proposed model. F1-Score is the weighted average of precision and recall. Accuracy is the number of suitably predicted data points out of all the data points.

	Precision	*Recall*	*F1-score*	*Support*
Class 0	0.94	0.87	0.90	468
Class 1	0.00	0.00	0.00	0
Class 2	0.00	0.00	0.00	0
Class 3	0.87	0.85	0.85	416
Class 4	0.00	0.00	0.00	0
Class 5	0.66	1.00	0.67	156
Micro average	0.87	0.87	0.87	1040
Macro average	0.38	0.46	0.40	1040
Weighted average	0.89	0.87	0.88	1040
Accuracy is	0.876935			

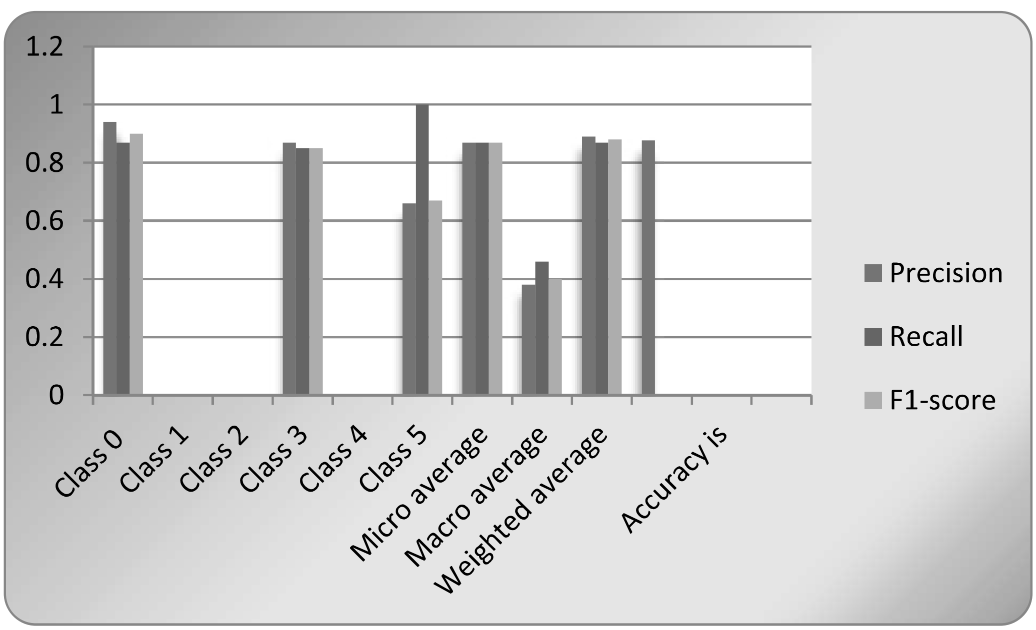

Figure-2: Graphical representation of the proposed Method (Ensemble learning method)

Table 2 provides the performance outcomes of the suggested algorithm. A bar chart, a data visualisation tool, is used in Figure 2 to graphically display the results of Table 2 results. The suggested method divides soil into six categories: class 0 is red soil, class 1 is black soil, class 2 is laterite soil, class 3 is sandy soil, class 4 is clay soil, and class 5 is loam soil (Alluvial Soil). Our suggested approach identifies three soil types, namely Class 0 (Red Soil), Class 3 (Sandy Soil), and Class 5(Loam Soil), out of a total of six soil types (Alluvial Soil). Our suggested approach has a predicted accuracy of 87% overall. The aforementioned Table-2 also calculates and records the weighted, micro, and macro averages.

Comparative Analysis of Existing Techniques with Proposed Model (Ensemble learning method)

With the aid of the findings provided in Table-3 below, we are comparing the outcomes received from the existing algorithms with the proposed algorithm. For easier comprehension, the comparison is also represented visually in the below figure-3 as a bar chart.

Algorithms	Precision	Recall	F1-Score	Accuracy
KNN	0.82	0.83	0.84	0.8356872
SVM	0.81	0.82	0.81	0.802529
DECISION TREE	0.79	0.81	0.7805	0.794765
Bayesian approach to Decision Tree	0.86	0.85	0.84521	0.840508
Proposed Algorithm	0.89	0.87	0.88	0.876953

Table-3: Comparative Analysis of the existing algorithms with the proposed model (Ensemble learning method)

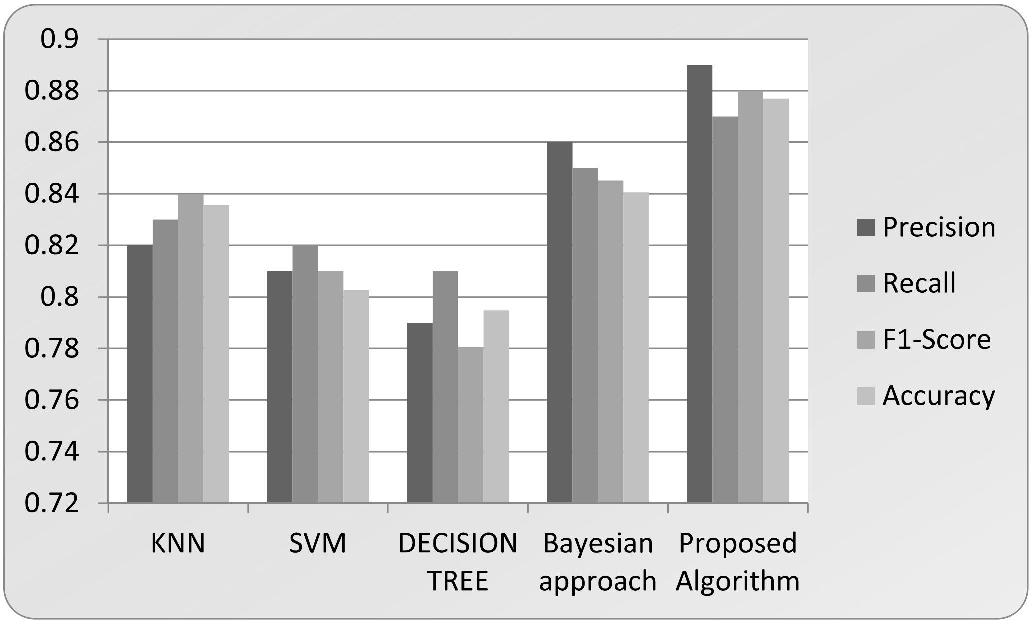

Figure-3: Comparisons of existing algorithms with the proposed algorithm (Ensemble learning method).

The suggested method has been compared to existing algorithms like Bayesian Approach to the Decision Tree model, KNN, SVM, Decision Tree, and other existing algorithms. The findings from both existing and proposed models are then documented in the table above. Figure 3 above uses a bar chart to visually represent the findings of the comparisons. So, it is clear why the suggested algorithm produces greater precision, recall, F1-score, and accuracy than the techniques currently in use.

Conclusion

The application of data mining techniques in the agriculture industry, specifically in soil type classification, presents a promising avenue for improving crop recommendations and enhancing agricultural practices. In this study, we addressed the critical issue of precise soil type categorization for a specific region, providing valuable guidance to farmers in selecting the most suitable crops. To achieve accurate and effective soil type prediction, we employed the Stacking approach of Ensemble Learning in conjunction with a Decision Tree model. This method demonstrated superior performance when compared to various classification algorithms, including K-Nearest Neighbor (KNN), Support Vector Machine (SVM), traditional Decision Tree, and a Bayesian approach to the

Decision Tree Algorithm. In the result and discussion section of this paper, we presented a comprehensive analysis of the soil type classification using these algorithms. To enhance clarity and facilitate understanding, we visualized the results through a ba-chart, providing a clear comparison of the classification performance. Notably, our findings strongly support the superiority of the proposed algorithm over existing methods. It consistently achieved better results, signifying its potential to significantly impact agricultural decision-making and contribute to improved crop yields. The proposed model's accuracy and effectiveness in soil type classification make it a valuable tool for assisting farmers in making informed decisions about crop selection based on soil characteristics. In conclusion, this research underscores the importance of leveraging advanced data mining techniques to enhance agricultural practices. The proposed Stacking Ensemble Learning approach with the Decision Tree model represents a valuable contribution to the field, offering a more accurate and reliable method for soil type classification, ultimately benefiting farmers and the agricultural industry.

References

[1]. Alobaidi. M.H, Meguid M.A, and Chebana, Ensemble learning for classification of Soil Liquefactions, CSCS Annual Conference, June 12-15, 2019.

[2]. Ramesh Babu Palepu and Rajesh Reddy Muley, An Analysis of Agricultural Soil by using Data Mining Technique International Journal of Engineering Science and Computing, Volume:7, Issue No:10, pp.15167-15177, October 2017, ISSN:2321-3361.

[3]. P.Surya and Dr.I.Laurence Aroquiaraj, Crop yield prediction in Agriculture using Data Mining Predictive Analytic Techniques International Journal of Research and Analytical Review (IJRAR), vol.5, Issue 4, pp. 783-787, December 2018, ISSN 2348- 1269.

[4]. Vrushal Milan Dolas, and Prof.Uday Joshi, A Novel Approach for Classification of Soil and Crop Prediction, International Journal of Computer Science and Mobile

Computing (IJCSMC), Vol.7, Issue 3,pp.20-24, March 2018, ISSN 2320-0887.

[5]. Rushika Gadge, Juilee Kulkarni, Pooja More, Sachee Nene, and Priya. R, Prediction of Crop yield using Machine-learning, International Journal Research Journal of Engineering and Technology (IRJET), volume: 05, issue: 02, pp. 2237-2239, Feb 2018, e-ISSN: 2395-0056.6.

[6]. N.Saranya and A.Mythili, Classification of Soil and Crop Suggestion using Machine-learning Technique, International Journal of Engineering Research & Technology (IJERT),Vol. 9 Issue 02, pp. 671-673 February-2020, ISSN: 2278-0181.

[7]. Chandan1, and Ritula Thakur, Recent Trends of Machine-learning in Soil Classification: A Review, International Journal of Computational Engineering Research (IJCER), vol: 08, Issue. 9, pp. 25-31, September – 2018, ISSN (e): 2250 – 3005.

[8]. Dr. D. Ashok Kumar, and N. Kannathasan, A Survey on Data Mining and Pattern Recognition Techniques for Soil Data Mining, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 1, pp. 422-426, May 2011, ISSN (Online): 1694-0814.

[9]. Ashwini Rao, Janhavi V, Abhishek Gowda N. S, and Mrs.Rafega Beham, Machine-learning in Soil Classification and crop detection, International Journal of Scientific Research & Development, Volume:4, Issue: 1, pp.792-794, April 2016, ISSN (online): 2321-0613".

[10]. P. A. Harlianto, T. B. Adji and N. A. Setiawan, "Comparison of machine-learning algorithms for soil type classification," 2017 3rd International Conference on Science and Technology - Computer (ICST), Yogyakarta, Indonesia, 2017, pp. 7-10, doi: 10.1109/ICSTC.2017.8011843.

[11]. N. Ajithkumar, P. Aswathi and R. R. Bhavani, "Identification of an effective learning approach to landmine detection," 2017 1st International Conference on Electronics, Materials Engineering and Nanotechnology (IEMENTech), Kolkata, India, 2017, pp. 1-5, doi: 10.1109/IEMENTECH.2017.8077018.

[12]. S. Raskar, A. Abhang, N. Pethe and V. Attar, "Crop Yield Prediction and Recommendation System," 2022 IEEE Pune Section International Conference (PuneCon), Pune, India, 2022, pp. 1-6, doi: 10.1109/PuneCon55413.2022.10014957.

[13]. Y. Garg, A. Masih and U. Sharma, "Predicting Bridge Damage During Earthquake Using Machine-learning Algorithms," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 725-728, doi: 10.1109/Confluence51648.2021.9377100.

[14]. S. Bhansali, P. Shah, J. Shah, P. Vyas and P. Thakre, "Healthy Harvest: Crop Prediction and Disease Detection System," 2022 IEEE 7th International conference for Convergence in Technology (I2CT), Mumbai, India, 2022, pp. 1-5, doi: 10.1109/I2CT54291.2022.9825446.

[15]. R. Kumar, N. Shukla and Princee, "Plant Disease Detection and Crop Recommendation Using CNN and Machine-learning," 2022 International Mobile and Embedded Technology Conference (MECON), Noida, India, 2022, pp. 168-172, doi: 10.1109/MECON53876.2022.9752173.