iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** VII **Month of publication:** July 2020

DOI: https://doi.org/10.22214/ijraset.2020.30611

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

Soil Classification and Crop Suggestion using Machine Learning

Archana Reddy K M[1], Chithra S[2], Hemashree HM[3], Mrs. Thanu Kurian[4]
[1, 2, 3]*Undergraduate students,* [4]*Assistant prof., Department of Computer Science, MVJCE*

Abstract: ***Agriculture is the basic source of food supply in all the countries of the world—whether underdeveloped, developing or developed. Not only providing food, this sector has contributions to almost every other sector of a country. According to the Bureau of Statistics (BS), 2017, about 17 % of the country's Gross Domestic Product (GDP) is a contribution of the agricultural sector, and it employs more than 45% of the total labour force. In the decreasing crop production and shortage of food across the world, one of the crucial criteria of agriculture now-a-days is selecting the right crop for the right piece of land at the right time. Hence, in our research we have proposed a method which would help suggest the most suitable crops for a specific land based on the analysis of the data of previous soil series classification using machine learning. In our work, we have implemented Bagging Classifier, Support Vector Machine, and k-Nearest Neighbour for soil classification and crop recommendation. We have trained these algorithms with the training data and later these were tested with test dataset. Then compared the performances of all the tested methods to arrive at the best outcome.***
Keywords: Machine Learning Algorithms, Artificial Neural Network, Crop Selection.

I. INTRODUCTIÓN

For a country, one of the most crucial aspects of its development circles around its capacity to produce food. From many years, agriculture has been associated with the production of essential food crops. The rate of urbanization at present is by-far the most superior aim of our civilization. In doing this, we are ignorantly diminishing our capacity for agriculture; especially in terms of land and fertility. As the amount of land will not be increasing in this era of urbanization and globalization, we will have to focus on making the most of what we have. Machine learning is used worldwide these days with great efficiency in sectors such as forecasting, pattern recognition, fraud/fault detection, prediction, virtual assistance, image processing, robotics, artificial intelligence and so on and so forth. The agricultural sector has work done on it through weather forecasting, crop disease prediction, yield prediction, etc. However, what we propose, which is soil classification and crop prediction itself, has not yet been implemented on a mass scale over large datasets and establishing a relationship among them to solve the problem through data analysis [1]. Introduction of data mining in agricultural field has made benefits in research field. Classification is very important in any field of science to establish the fundamentals. It helps to find the diversity between the objects and concepts. It also provides necessary information through which research can be made in a systematic manner [2]. Soil is one of the key components in agricultural field for yielding crops. Soil classification philosophies follow the existence knowledge and practical circumstances. On the land surfaces of earth, classification of soil creates a link between soil samples and various kinds of natural entity [3] Machine learning is the branch of computer science which is used to build algorithms which exhibit self-learning i.e. learning which is done by the machine itself hence the "Machine Learning". To show intelligence machine needs to interpret and analyse the input. After analysing it the result data apart from simply following the instructions on that data. This is the thing that machine learning algorithms do. Machine learning focus on the development of computer programs that can get data and utilize it to learn. The way toward learning starts with perceptions on information, for example, direct experience, or instruction, in order to look for patterns in data. It helps to make better decisions in the future based on the examples that we give. The essential point is to permit the computer learn automatically without human intercession or help and regulate actions consequently. It is a major field in computer science which is being utilized in various forms of progressive technological development programs all around the world. It is regarded as the future of engineering and Artificial Intelligence [13]

Soil is composed of organic matter, gases, minerals, liquids and organisms that provides substratum for plant and animal life. Soil classification plays critical role in crop management, soil improvement, land consolidation, management of drainage, soil erosion and irrigation. The physical and chemical parameters measured from real-time field samples are the most influential features in soil characterization. Physical parameters such as moisture and temperature are related to the organization of the particles and pores, reflecting effects on root growth, speed of plant emergence and water infiltration.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

II. PROPOSED METHODOLOGY

The System architecture of the proposed model is shown in fig. 1.

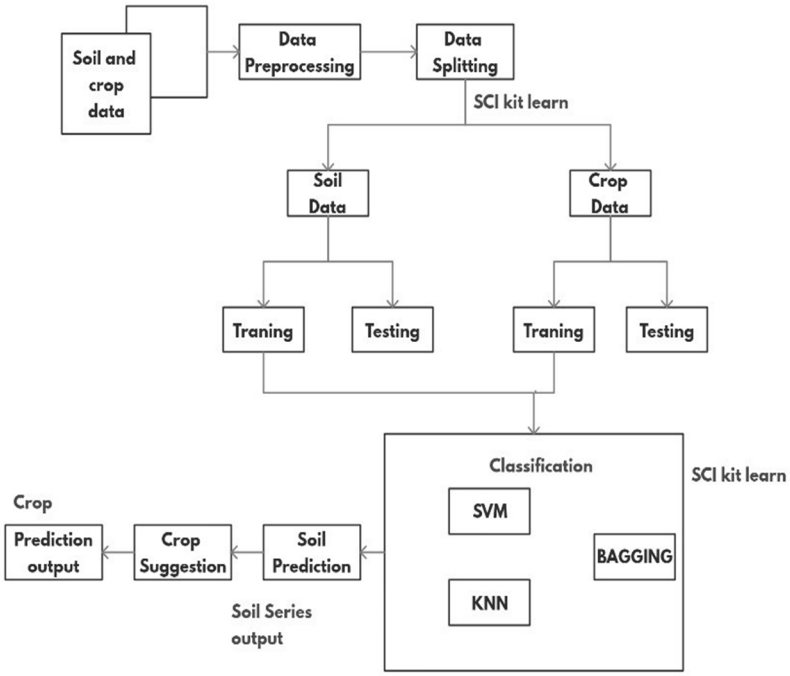


Fig. 1: Proposed System Architecture

The proposed method involves two prediction schemes. They are soil classification and crop suggestion. Two datasets are used: Soil dataset and crop dataset. Soil dataset contains class labelled chemical features of soil. The crop suggestion dataset contains class labelled crop suggestion attributes. The details of the datasets are listed in the table below.

Table I: soil classification attributes:

Attributes	Details
Ca	Calcium Content of the Soil
Mg	Magnesium Content of the soil
K	Potassium Content of the soil
S	Sulphur Content of the soil
N	Nitrogen Content of the Soil
Lime	Lime Content of the Soil
C	Carbon Content of the Soil
P	Phosphorous Content of the Soil
Moisture	Moisture level of the soil

Table II: Crop Suggestion Attributes:

Attributes	Details
Soil	Soil Series
Temperature	Temperature Geographical Data
Humidity	Humidity Geographical Data
ph.	Ph Geographical Data
Rainfall	Locality Rainfall Level

A. *Data Analysis*

The first step in the proposed system is the data analysis. The collected the soil and crop related data are analysed manually for finding the duplication and null values.

B. *Data Pre-Processing*

After successful data analysis, the data are subjected to pre-processing as pandas data frame. The data frame values are replaced with nominal values if there are any null values using NumPy feature. The other important aspect done here is checking the duplicate number of rows in the features and are skipped if present.

C. *Data Splitting*

The third important step in the proposed system is data splitting. The original set of features and labels in both soil and crop data are split into training and testing set using scikit model selection function which results in training set of features and labels as well as the testing set of features and labels.

D. *Data Preparation*

The split data is then subjected to preparation by standardization and sampling procedure for the minority and majority label class imbalance problem.

E. *Classification*

This is final step in the proposed system. The classification is performed in two steps. The first approach is the soil classification. The user will feed in the chemical attributes of the soil which results in the soil series prediction output. The predicted soil series output is given as input for the crop suggestion which recommend the corresponding crop for the living environment. The classifier used in the proposed system are K-Nearest Neighbour, Bagging and Support Vector Machine classifier.

III. CONCLUSION

Agriculture is a fundamental aspect of modern civilization. With increasing world hunger and economy breakdown, the proper selection of crop emerges as a massive factor in this. Our proposed model can predict the proper soil series and crop for a particular piece of land in a way that is very efficient. We have implemented 3 different types of machine learning in this research. All the accuracy of the models was carefully obtained through various different methods, and compared with each other. Using multiple algorithms helped to understand which algorithm is more suitable for this system. The crops can be predicted based on a very suitable set of features included in the dataset used. From this research work, we found that, the cleaner the data, the better the accuracy of the result. The entire length of this research was very enjoyable, as we were able to work in the field of machine learning and deep learning. Some python library usage, algorithm fitting, and accuracy checking methods were very interesting in practicality. We trust all our algorithms and research work to efficiently work on any platform and any new type of data. The predictions made were solid and robust. Such strength in the model delights us, and we hope this keeps working over the years without issues.

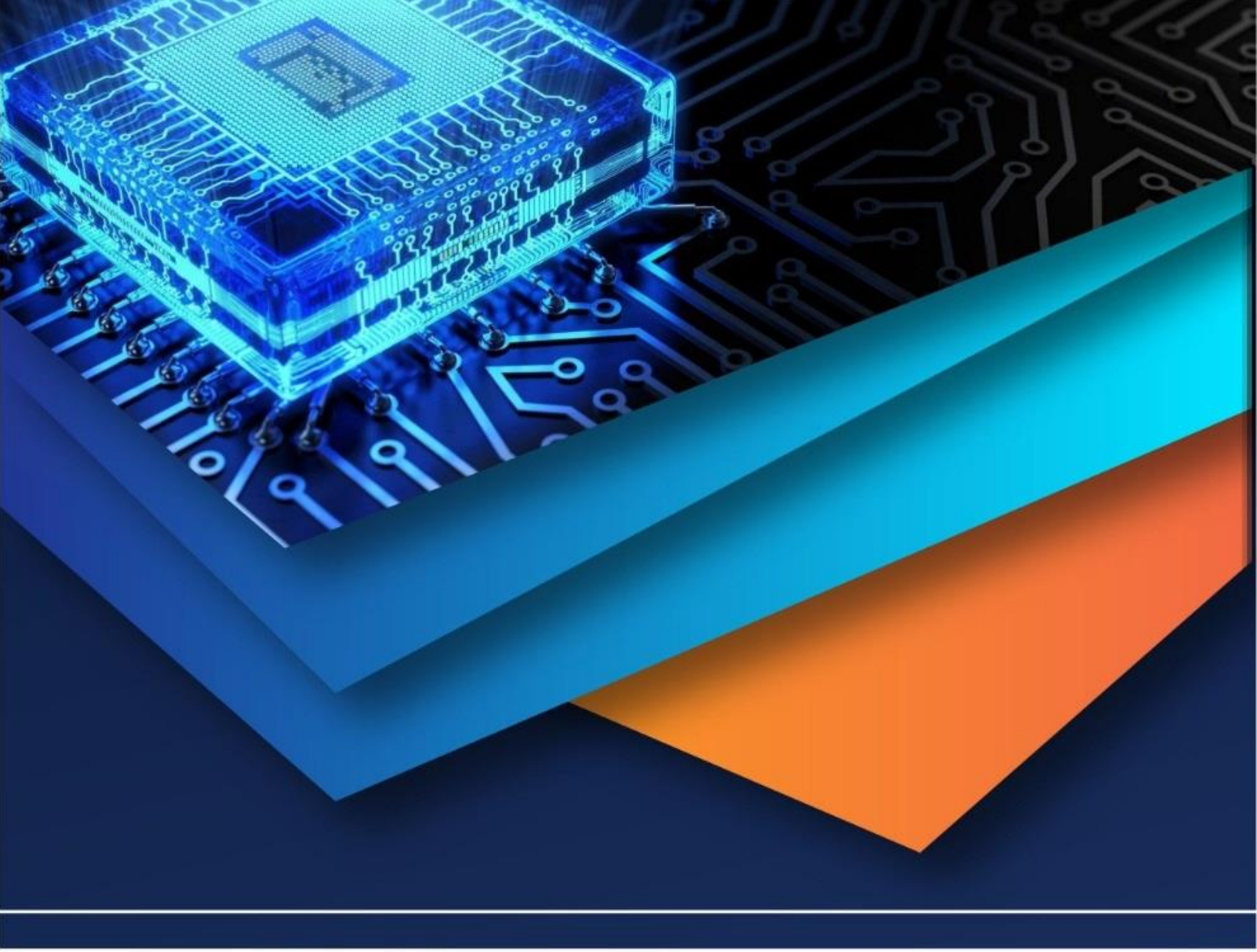
IV. FUTURE WORK

In the future, we hope that this model would be implemented with much more efficient dataset for a specific piece of land containing information such as different soil property, soil pH, different mineral percentage etc. So that no agricultural plot be wasted by harvesting a less efficient crop. We want this model to be used worldwide for the further development of the agricultural sector. This can be a field of interest for both entrepreneurs and researchers. In this research, we have developed a model that will predict suitable crop or predict the viability of a specific crop for a particular plot of land. Our plan is to build on this research and improve this model; mostly by adding further algorithms based on other aspects of machine learning. For future work, we also want to build a platform for all the farmers who will be using this model, to share the predictions on their land with other farmers all over the land. This will let a farmer in one country to know the prospect of farming in another part of the world; down to a specific geographical unit. The accuracy values we obtained in the specific crop prediction part of our research was very poor to our target standards. In the future, we hope to implement ANN for crop prediction and check its viability on this. In addition, we desire to take this model into the mobile phone platform. Android and iOS applications will be a part of it. Nowadays, even some the poorest farmers are seen to be using such devices. A mobile application can be implemented into help them as well. One of our most

ambitious goals is to improve our model so that it can give a string of crop sequencing predictions. This will enable us to derive predictions for not just the next year, but of multiple years ahead. Now we have only implemented supervised learning. In the future, we plan to use unsupervised and reinforced learning as well. That will give a new dynamic to our research. We are looking for a better world in a sense that we hope our model to be able to successfully raise the standard of farming and agriculture. We hope, whatever shapes this model take, it will stay user-friendly and properly welcomed by the potential users.

REFERENCES

[1] Aggarwal, C. C. (2014). Data classification: algorithms and applications. CRC Press.
[2] Breiman, L. (2001). Random forests. Machine learning, 45(1):5–32.
[3] Freund, Y., Schapire, R. E., et al. (1996). Experiments with a new boosting algorithm. In Icml, volume 96, pages 148–156. Citeseer.
[4] Hsu, C.-W., Chang, C.-C., Lin, C.-J., et al. (2003). A practical guide to support vector classification.
[5] Jain, N., Kumar, A., Garud, S., Pradhan, V., and Kulkarni, P. (2017). Crop selection method based on various environmental factors using machine learning. [6] Kaur, K. (2016). Machine learning: applications in indian agriculture. Int. J. Adv. Res. Comput. Commun. Eng.(IJARCCE), 5(4).
[6] Kesavaraj, G. and Sukumaran, S. (2013). A study on classification techniques in data mining. In Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on, pages 1–7. IEEE.
[7] Kim, S.-B., Han, K.-S., Rim, H.-C., and Myaeng, S. H. (2006). Some effective techniques for naive bayes text classification. IEEE transactions on knowledge and data engineering, 18(11):1457–1466.
[8] Kolesnikova, A., Song, C.-H., and Lee, W. D. (2009). Applying uchooboost algorithm in precision agriculture. In Proceedings of the International Conference on Advances in Computing, Communication and Control, pages 30–34. ACM.
[9] Kotsiantis, S. B., Zaharakis, I., and Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160:3–24.
[10] Kumar, R., Singh, M., Kumar, P., and Singh, J. (2015). Crop selection method to maximize crop yield rate using machine learning technique. In Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2015 International Conference on, pages 138–145. IEEE.
[11] Maind, S. B., Wankar, P., et al. (2014). Research paper on basic of artificial neural network. International Journal on Recent and Innovation Trends in Computing and Communication, 2(1):96–100.
[12] Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (2013). Machine learning: An artificial intelligence approach. Springer Science & Business Media.
[13] Osman, J., Inglada, J., and Dejoux, J.-F. (2015). Assessment of a markov logic model of crop rotations for early crop mapping. Computers and Electronics in Agriculture, 113:234–243.
[14] Rush, S. (2001). Logistic regression: The standard method of analysis in medical research. Technical report, Technical Report Mathematics.
[15] Su, Y.-x., Xu, H., and Yan, L.-j. (2017). Support vector machine-based open crop model (sbocm): Case of rice production in china. Saudi journal of biological sciences, 24(3):537–547.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY