EasyChair Preprint

№ 10175

Crop Selection According to Environment Using Latest Technique of Machine Learning

Rajkumar Yathirajyam, Mohit Malik, Manish Kumar,
Divyam Sharyan, Vipul Nuthalapati and Pravinkumar Tiwari

May 16, 2023

CROP SELECTION ACCORDING TO ENVIRONMENT USING LATEST TECHNIQUE OF MACHINE LEARNING

EFFICIENTNET

Rajkumar Yathirajyam
Department of computer Science and Engineering
Lovely Professional University
Phagwara, Punjab
rajkumaryathirajyam@gmail.com

Assistant Professor Mohit Malik
Department of computer Science and Engineering
Lovely Professional University
Phagwara, Punjab
mohit.29613@lpu.co.in

Manish Kumar
Department of computer Science and Engineering
Lovely Professional University
Phagwara, Punjab
manish215555@gmail.com

Divyam Sharyan
Department of computer Science and Engineering
Lovely Professional University
Phagwara, Punjab
Sharyandivyam54@gmail.com

Nuthalapati Vipul
Department of computer Science and Engineering
Lovely Professional University
Phagwara, Punjab
vipulchowdary1@gmail.com

Pravin Kumar Tiwari
Department of computer Science and Engineering
Lovely Professional University
Phagwara, Punjab
tiwaripravin114@gmail.com

ABSTRACT: The study of agriculture is gaining popularity. In agriculture, crop prediction is very important because soil factors like temperature, humidity, and rainfall have a big effect. Ranchers used to have the option to pick the yield they needed to plant, screen its encouraging, and choose when to reap it. Notwithstanding, the quick changes in the climate have made it unthinkable for farmers to do as such. Thusly, lately, machine learning calculations play assumed the part of forecast, and in this review, farming creation was determined utilizing different these methodologies. To guarantee the accuracy of a particular machine learning (ML) model, effective feature selection methods must be used to preprocess raw data into a dataset suitable for machine learning. Only data characteristics that have a significant impact on the model's output should be included in order to reduce duplication and improve model quality. The model only has the most important features because the best feature selection was made. If every characteristic from the raw data is joined without their value during the model-building process, our model will be too complicated. In addition, the model's output accuracy would be decreased if factors that have little effect on the model were included.

Keywords – <*Agriculture*>, <*Crop forecast*>, <*Feature selection*>.

1. INTRODUCTION

In agriculture, crop prediction is a complicated process that requires a variety of suggested and tested models. The problem necessitates the use of a variety of datasets because agricultural production is influenced by both biotic and abiotic variables. Bibiotic variables, which are elements of the environment, are produced when living things like microbes, plants, animals, parasites, predators, and pests interact with one another. This category also includes human-caused factors like fertilizer, irrigation, plant protection, air and water pollution, soil contamination, and so on. These yield creation difficulties might bring about plant yield varieties in synthetic organization, inner blames, and structure issues. Both biotic and abiotic factors influence plant growth and quality, as well as the production of the environment. Substance, physical, and other abiotic factors are recognized. Mechanical vibrations (vibration, clamor) and radiation, (for example, ionizing, electromagnetic, bright, and infrared) are instances of actual factors. environment (temperature, mugginess, development of air, and daylight); landscape, the dirt's roughness, the climate, and the kind of soil; science of water, especially saltiness. Sulfur dioxide and its subsidiaries, PAHs, nitrogen oxides and their subordinates, fluorine and the mixtures it contains, lead and the mixtures it contains, cadmium and the mixtures it contains, nitrogen manures, pesticides, and carbon monoxide are instances of compound parts. Asbestos, mercury, arsenic, dioxins and furans, and aflatoxins are extra synthetic substances. Its qualities are affected by abiotic factors like bedrock, relief, climate, and water conditions. There are many ways that soil-forming forces affect soil formation and agricultural value.

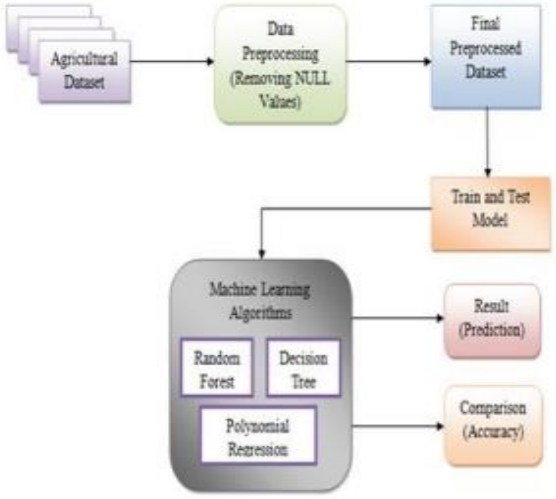


Fig.1: Example figure

2. LITERATURE REVIEW

Applying naive Bayes classification technique for classification of improved agricultural land soils:

PC and information stockpiling headways have brought about monstrous measures of information. New technologies and methods, such as information mining, have been developed to help close the information gap. The difficulty has been figuring out how to extract data from this crude data. This study set off to look at these noteworthy data mining strategies and apply them to a dirt science informational index to check whether any critical associations

could be laid out. A vast selection of soil data sets has been made available by the Division of Soil Sciences and Rural Science at the S.V. Farming School.

Biotic components influencing the yield and quality of potato tubers

Over the past ten years, potato yields in Canterbury have remained roughly constant at 60 t/ha. On the other hand, yields of up to 90 t/ha are predicted by potato development models, which some commercial growers have already achieved. A two-year study on farming production constraints was led by partners from industry and academia. For the essential laying out season, 11 dealing with crops were steadily chipped away at.

Response surface methodology: A retrospective and literature survey

RSM is a strategy for additional upgrading cycles and thing plans that joins real preparation and numerical smoothing out methods. The first studies in this field were done in the 1950s, and they have been used a lot, especially in the cycle and substance industries. In the past 15 years, RSM has seen a lot of use and some amazing advancements.

Application of response surface methodology for optimization of potato tuber yield

The creator researches the urgent functional parts of augmenting potato tuber yield in Kenya in this work. Potato producers will need to avoid adding more information along these lines. Factorial plans 2 and 3, as well as the response surface strategy, were used to extra the potato creation process. Utilizing a reaction surface methodology, the consolidated impacts of water, nitrogen, and phosphorus mineral upgrades were explored and created. The best creation limits for potato tuber creation were viewed as 70.04% water structure water, 124.75 kg/ha urea nitrogen, and 191.04 kg/ha triple super phosphate phosphorus. A potato tuber yield of 19.36 kg per plot estimating 1.8 meters by 2.25 meters can be accomplished under ideal circumstances.

Improving potato yield prediction by combining cultivar information and UAV remote sensing data using machine learning

Accurate high-goal yield guidelines are used in accuracy agriculture to identify global yield inconsistency designs, describe essential components that contribute to yield changeability, and provide information about site-specific management. Changes to cultivars in guaging potato (Solanum tuberosum L.) might have an effect on the production of tubers by using technologies like remote sensing. By combining cultivar data with machine learning (ML) calculations that made use of remote detection by unmanned aerial vehicles (UAVs), this study aimed to increase potato production gauge. All through the improvement season, various regular photos of UAVs were taken.

Machine learning (ML) algorithms Random Forest Regression (RFR) and Support Vector Regression (SVR) were utilized to associate cultivar data and different vegetative focuses. It was found that horrible information accumulated by UAVs at the tuber start stage in the early developing season (late June) had a more grounded connection with potato charming yield than horrendous information accumulated later in the developing season.

3. METHODOLOGY

Calculating the agroclimatic parameters that influence the growth of winter plant species, particularly grains, in the cool temperature zone is the primary test. Wintering yield is altogether affected by the aggregate and recurrence of days with temperatures over 5 degrees Celsius, as well as the quantity of days with temperatures somewhere in the range of 0 and 5 degrees Celsius. Using publicly available data, many of these can be examined, resulting in years-long relapse measures. The need for a state strategy of grain market mediation has been analyzed using models that have been developed. For valid creation hypotheses to be formed, meteorological boundary estimation is essential.

Disadvantages:

1. Soil and meteorological factors like rainfall, humidity, and temperature have a significant impact on crop prediction in agriculture.

2. Due to rapid changes in the environment, farmers have been unable to continue farming.

There are numerous issues in this area of research. The results of crop prediction models are currently satisfactory, but they may be better.

Advantages:

1. To avoid obvious repetition and increase ML model accuracy, only information credits with a high degree of value in determining the model's final outcome should be included.

2. In terms of prediction accuracy, an ensemble technique beats the previous classification strategy.

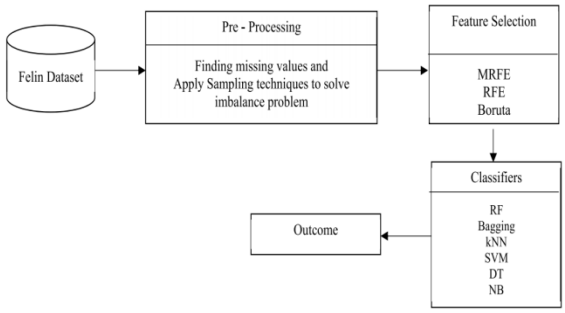


Fig.2: System architecture

MODULES:

We fostered the modules recorded underneath to finish the previously mentioned project.

- Exploration of data: Data will be entered into the system with the help of this module.

- Data will be read and processed by this module.
- Information will be separated into train and test fragments with this module.
- The most common way of building a model with and without include determination is known as model age. - SVM, Naive Bayes, KNN, Random Forest Decision Tree, and Voting Classifier. It was determined how accurate the algorithm was.
- Login and registration for users: You must first register and then log in to use this module.
- The use of this module will give projected yield.
- Prediction: a definitive anticipated esteem is shown.

4. IMPLEMENTATION

ALGORITHMS:

KNN: K-Nearest Neighbor is shortened to the acronym "KNN." It is a ML-managed computation. Relapse and grouping of problem articulations can both be handled by the strategy. The number of variables with closest neighbors that can be predicted or categorized is denoted by the letter "K."

Naive Bayes: Naive Bayes characterisation is a probabilistic classifier. Probability models with high freedom suspicions are used to establish it.

Random forest: It might be used to deal with ML issues like planning and relapse. It is based on the idea of collecting learning, in which multiple classifiers are combined to work on the model's presentation and solve a difficult problem. As its name suggests, Random Forest is a classifier that "works on the extended exactness of that dataset" by "taking the normal of several decision trees on diverse subsets of the supplied dataset."

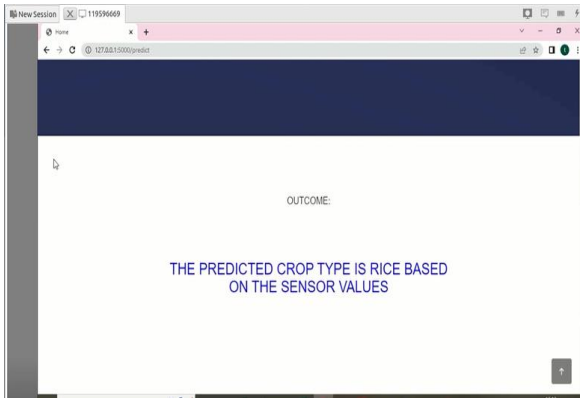
Decision Tree: While choosing whether or not to part a hub into at least two sub-hubs, decision trees utilize various techniques. The consistency of the subnodes is improved by their appearance.

SVM: The Support Vector Machine (SVM) is a notable Overseen Learning approach for Portrayal and Backslide. Nonetheless, most of its ML applications manage characterisation issues. The objective of the SVM technique is to find the best line or decision limit for n-layered space design so new data can be effortlessly added to the right characterization later. The cutoff that is most sensible is a hyperplane.

Voting classifier: A machine learning estimator known as a voting classifier is capable of predicting the outcomes of numerous base models or estimators. Each estimator output's

aggregate criterion could be paired voting decisions.

5. EXPERIMENTAL RESULTS



6. CONCLUSION

In agriculture, it is difficult to predict development crops. Plant development yield size was estimated using a variety of element determination and order methods in this study. The results suggest that a group method performs better than the existing order technique in terms of expectation accuracy. If ranchers and nations know where to harvest potatoes, cereals, and other energy sources, they may find it easier to organize their planting efforts. Techniques for modern estimation may lead to significant financial gains.

REFERENCES

[1] R. Jahan, ‘‘Applying naive Bayes classification technique for classification of improved agricultural land soils,’’ Int. J. Res. Appl. Sci. Eng. Technol., vol. 6, no. 5, pp. 189–193, May 2018.

[2] B. B. Sawicka and B. Krochmal-Marczak, ‘‘Biotic components influencing the yield and quality of potato tubers,’’ Herbalism, vol. 1, no. 3, pp. 125–136, 2017.

[3] B. Sawicka, A. H. Noaema, and A. Gáowacka, ‘‘The predicting the size of the potato acreage as a raw material for bioethanol production,’’ in Alternative Energy Sources, B. Zdunek, M. Olszáwka, Eds. Lublin, Poland: Wydawnictwo Naukowe TYGIEL, 2016, pp. 158–172.

[4] B. Sawicka, A. H. Noaema, T. S. Hameed, and B. Krochmal-Marczak, ‘‘Biotic and abiotic factors influencing on the environment and growth of plants,’’ (in Polish), in Proc. Bioróżnorodność Środowiska Znaczenie, Problemy, Wyzwania. Materiały Konferencyjne, Puławy, May 2017. [Online]. Available:

[5] R. H. Myers, D. C. Montgomery, G. G. Vining, C. M. Borror, and S. M. Kowalski, ‘‘Response surface methodology: A retrospective and literature survey,’’ J. Qual. Technol., vol. 36, no. 1, pp. 53–77, Jan. 2004.

[6] D. K. Muriithi, ‘‘Application of response surface methodology for optimization of potato tuber yield,’’ Amer. J. Theor. Appl. Statist., vol.

4, no. 4, pp. 300–304, 2015, doi: 10.11648/j.ajtas.20150404.20.

[7] J. R. Olędzki, ‘‘The report on the state of remotesensing in Poland in 2011–2014,’’, Remote Sens. Environ., no. 2, pp. 113–174, 2015.