Sell Quality determination for Chemical Fertilizer Data using Expert System

Mahendra Kumar Ratre[a], Dr.Ravindra Kumar Gupta[b], Dr. Shikha Agrawal Associate[c]

[a]Research Scholar, Deptt. of C.S.E, SRK University, Bhopal, India
[b]Department of Computer Science & Engineering, SRK University, Bhopal, India
[c]Professor, Department of Computer Science & Engineering, UIT- Rajiv Gandhi Proudyogiki Vishwavidyalaya Bhopal, India
E-mail: [a]mahendraratre@gmail.com,[b] ravindra_p84@rediffmail.com,[c] shikha@rgtu.net

Article History Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: There is an growing appetite for food due to the ever increasing global population, so new technologies need to be created to increase crop yield. This paper proposes an intelligent way of forecasting crop yield and recommends the best variables for optimising crop yield. With technical developments, the emphasis has now moved to the use of computers and control systems for process management and efficiency enhancement. we estimate the crop yield per acre, in this research work to proposed hybrid approach for Chemical Fertilizer Data classification using SVM and neural network approach with expert system improvement. Yield and data obtained from Madhya Pradesh of Agriculture are used in the proposed process. Humidity, yield, temperature and rainfall are the different parameters used in the dataset.

Keywords: Support vector machine, artificial neural network, Crop Yield Production, Machine Learning.

1. Introduction

The key purpose of agricultural planning is to attain the optimum rate of crop production through the use of a finite number of soil resources. Many machine learning algorithms may help to improve the rate of crop yield production. We can apply crop selecting technique and decrease the losses whenever there is loss in unfavourable conditions. And it can be used in favourable circumstances to increase crop yield rates. This maximisation of the rate of yield helps in the improving countries economy. We have some of the variables affecting the yield rate of crops. They are seed efficiency and crop selection. Before sowing, we need to test the consistency of the seeds. As we know, good seed quality allows to produce a greater rate of production. And crop selection relies on two factors, which are favourable and unfavourable circumstances. By using techniques of hybridization, this can also be strengthened. Many surveys are being undertaken to enhance agricultural preparation. The aim is to get the full crop yield. Machine learning techniques can be used to improve the yield rate of crops. In order to increase the yield rate of crops, machine learning methods may be used. To boost crop production, the crop selection method is applied. Development of crops can depend on the region's geographical conditions, such as river land, hill areas or areas of depth. Climate conditions, such as humidity, snow, temperature, clouds. For different crops, several criteria are used to make various predictions. using analysis, these prediction models can be tested. Two ways are known as these predictions. One is the standard method of analytics and the other is the methods of machine learning. In forecasting single sample spaces, the conventional approach aids. And the techniques of machine learning help to model different predictions.

2. Problem Formulation

Several climatic conditions impact the development of agriculture. Like metrological parameters (humidity, wind speed, temperature and humidity), parameters of precipitation (rainfall, rainfall in the area, irrigation, etc.) and parameters of the soil (PH, organic carbon, phosphorus, fibre, etc.) And everything is messed up due to constant climate change. Farmers in India continue to pursue the old technologies they adapted from their ancestors. But the issue is that all was going on time was really safe in the earliest time setting. Yet today , thanks to global warming and many other causes, most aspects have changed. The biggest challenge with farming in India is the lack of seasonal rainfall. Humidity is also important for crops, but it has become excessive, as a downside it also converts. It has affected the winter season, so Rabi crops are widely affected. The rainfall in the winter season was as high as anticipated for a few years. We need to build a method that would be able to identify the secret facts or outcomes, trends and observations to solve these problems above. The farmer will foresee which crop should be sown in order to benefit more from it. We are applying data analytics techniques to agricultural production-based datasets in the proposed framework and finding insights to support farmers and their decision-making. We are proposing (SVM+ANN) a method based on descriptive analytics in this research article. That helps farmers to consider what has happened in the past and what is going to happen. So here we gather different data from the output of crops, rainfall and soil data and prepare their respective datasets. In this strategy, we use SVM+ANN to train the model and it will provide the approximate cost of generating crop yields and the corresponding output class.

3. Related Work

In this section we covert number work done by the different research work. We study and analysis machine learning and neural network algorithm for data analysis in Chemical Fertilizer Data using Expert System

Filippi, P. et ak[1] A tentative, data-driven approach to forecasting the yield of wheat, barley, and canola crops was proposed in this report, using only on-farm data and publicly available external data. The findings of this method are positive, and its generalised nature allows it to be extended to many other agricultural systems in which data on yields are available. Particular beneft was identified from the inclusion of field prior yield information.

Shah, A. et al[2] The system suggested uses yield and weather data obtained from the Agriculture department of the United States. Humidity, yield, temperature and rainfall are the different parameters used in the dataset.

[3] For maximum crop yields, this forecast can help farmers pick the most suitable temperature and moisture content.

Medar, R. et al[4] Using machine learning methods that are conveniently applicable to the farming industry, we will boost agriculture. Useful and reliable knowledge on various problems still plays an important role in this, along with all the developments in machinery and technology used in agriculture.

Meeradevi et al[5] ame regions , The proposed model, ame areas, is used to forecast crop yield. The goal of the proposed model is to create a method designed to provide forecasting based on the individual 's crop. The product edition being created relies on the forecast, but if more detail is available and used as features in the data model, it will potentially be included. The proposed model provides farmers with a robust collection of recommendations to maximise their crop selection based on individual variables such as location, farm scale, temperature , precipitation, and different datasets of crops.

Kumar, R et al[6] A system called CSM to pick the series of crops to be planted throughout the season is described in this article. The CSM approach will increase the seasonal net yield rate of the crops to be planted. The proposed approach resolves the selection of crop(s) based on parameter-influenced predictive yield rates (e.g. temperature, soil condition, water density, type of crop).

Vagh, Y. et al[7] agricultural context for the collection and interpretation of regional data on land use. The geographic data consisted of land use profiles for grain and cereal production. This were connected to previously reported climatic data from Australian fixed weather stations that were interpolated to match a surface grid using ordinary krigeing.

4. Proposed Methodology

For its economic growth, an agro-based country depends on agriculture. If the population of the world rises, the depend on agriculture also increases and the resulting economic development of the nation is impacted.In this case, the rate of crop yields plays an important role in the country's economic development. So, there is a need to boost the rate of crop yield. To solve this problem, certain biological approaches (e.g. seed quality of crops , crop hybridisation) and certain chemical approaches (e.g. fertiliser, urea, potash use) are carried out. In addition to these methods, a technique for seed arrangement are need to increase the net crop yield rate over the season. In order to design a smart data mining strategy with machine learning algorithms, crop yield output prediction. Achieving the net crop yield rate over the season. The crop can be defined as:

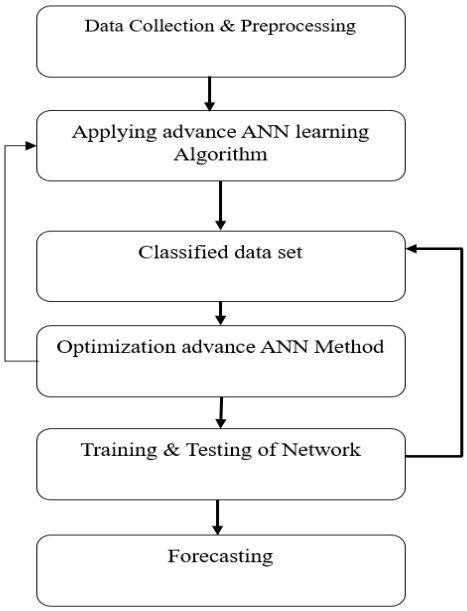


Figure 1: Data collection and classification process

5. Proposed algorithm

The whole season of the year in India depends on the depth of summer rainfall per year [5]. So, for the whole year, it is possible to forecast the season in advance. It is possible to forecast crop yield rates earlier on the basis of these expected seasons and past crop yield rate data. The SVM+ANN algorithm operates in advance on crop yield rate estimation based on desirable status and produces a series of crops with the highest net yield rate. Remember the seed sowing table, for instance,. Algorithm:

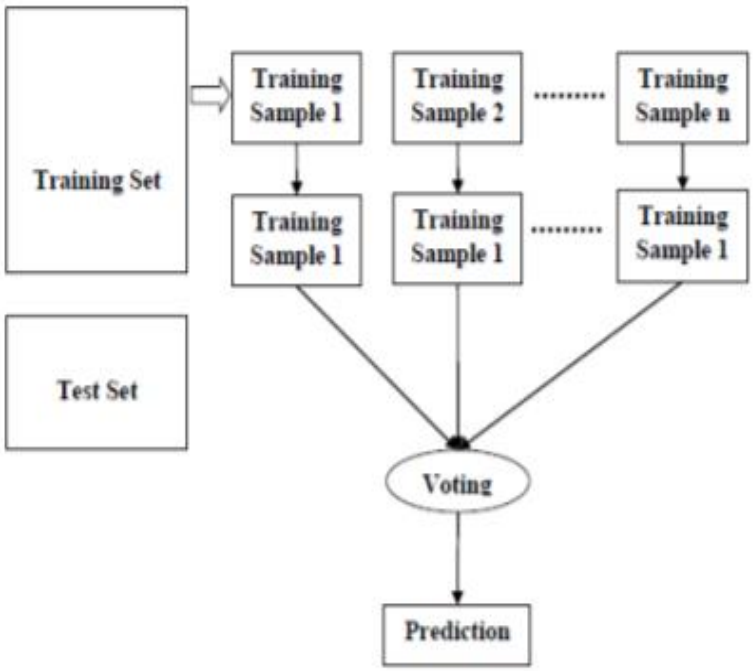
Proposed System: This condition focuses primarily on temperature calculation, forecasts of crop yields and anticipation of harvest costs. By recognising the natural segments, these elements allow the ranchers to produce the best nourishment harvests and raise the correct animals. Similarly, by changing planting periods, selecting assortments of various construction terms or modifying harvest pivots, ranchers can somewhat respond to shifts in the environment. For the research investigation, the factual numerical information is associated with horticulture is adopted. However, to deal with the observable knowledge obtained, the grouping-based schemes and the measurements administered are used. In comparison, for better order results, rational grouping methods such as Random Woodland (RF), Support Vector Machine (SVM) and Neural Networks are used. Algorithm Support Vector Machine (SVM). The measures for understanding SVM are as follows.

Step-1: Begin by determining arbitrary examples from the dataset in question.

Step 2 Next, for each case, this calculation will create an option tree. It will get the assumption out of each selection tree at that point.

Step 3 In this step, with each anticipated outcome, elective outcomes will be conducted.

Step 4-This is the last step in SVM, selecting as the final prediction outcome the most preparation and testing data collection for prediction



.

Fig. Shows the Working scenario of Support vector machine.

On large datasets, this equation runs proficiently and it has better grouping accuracy. Both programmes can help to predict precipitation, the determination of crop yield and the estimation of harvest costs. Exact evidence on harvest yield experience is crucial to deciding on choices found by executives' farming hazards. The paper suggests a plan to estimate the yield of the crop in this way. Until production on the farm, the rancher can verify the yield of the harvest according to the section of land. .

6. ANN+SVM

ANN+SVM is inspired model of natural neural system, which is made up of several interconnected neurons. The most efficient and utilized sort of ANN+SVM is the multilayer observation. The neurons on the yield layers are dynamic and the outcome is considered as the yield given by the system. There are some hidden layer between the data and yield layer. Every neuron can get input from the neuron has a place with past layer and it can send its yield to neurons that have a place with the progressive layer. BPNN is normally in light of the error back spread to the multi-layer Neural Network.

The principle ventures in Back Propagation calculation as takes after:
Step 1: Feed the standardized data to the system and compute the comparing comes about.
Step 2: Compare the error between ascertained result and actual results.
Step 3: The association weight and participation capacity is balanced in view of the error.
Step 4: If error more prominent than the resistance then go to step 1 else stop the Procedure.

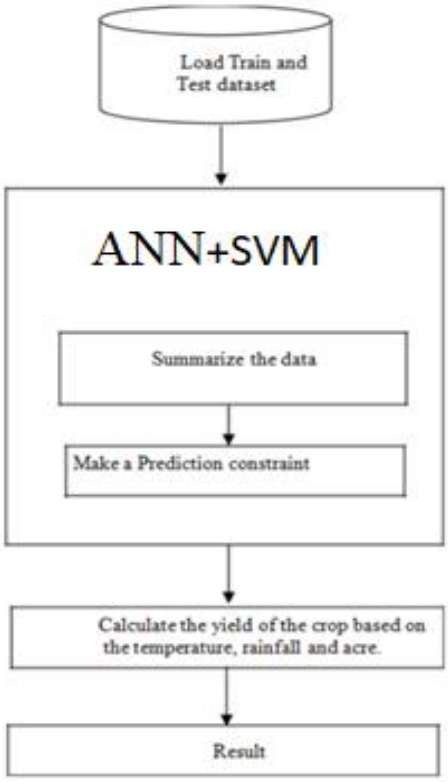


Figure 2: work flow ANN+SVM

We are going to use two distinct approaches here. First is the methodology of SVM and the method of Artificial Neural Network. By using these two techniques, we can obtain the consistency of results. A python application is generated in order to predict the crop yield rate. Three sections are included in this application. First, datasets are handled, second, datasets are checked, and third, datasets are evaluated. We can get the datasets of previous years when handling datasets and they can even be translated into a supporting format. Both datasets are translated to the attribute relationship file format as we are using the Weka tool in this project. We may do a single test in the research part. We've considered two machine learning approaches. One is ANN and the other method is SVM. In testing, we can pick any of the methods and conduct dataset testing, such as by choosing a specific crop, a specific position and a specific season, we can get yield results. We will enter a whole dataset file in the analysis section and obtain the precision of the two separate methods. It helps in predicting which technique is successful. In the agricultural sector, because farmers face many challenges, we need to mitigate their problems. This prevention of problems can be solved by the implementation of modern farming techniques. We should extend the techniques of machine learning to agriculture. We have strategies for clustering and grouping that can be used on crops. In order to increase the yield of crops, we may also add certain regression methods. We've only found the ANN+SVM approach in this project. We can forecast which crops will be picked for their soil and season using these two techniques. We also created a python framework because farmers don't know how to use the Weka programme. . This programme helps them estimate the yield. We may perform single tests here in this application by providing feedback such as crop name, season selected and position selected. With the ANN+SVM method, we can use either method. You can pick the form as soon as you enter the input and mine the output. The findings will inform you the rate of yield of that seed. And by evaluating the datasets, we can do different experiments. It helps you to pick a whole file at once while evaluating it to get the accuracy. We will do the different experiments explicitly here instead of trying to run single tests. This testing aims to achieve the consistency of two techniques. Through this, we will come to know which methodology among the methods given is fine. And this will assist farmers to pick the crop for their land or the area. The data sets contain data from the previous year's reports. Ses databases help to estimate the outcomes for new cases. Farmers will provide the test for every example and get the yield rate for the seed. Thus, this application allows farmers to choose the best crop for the land. And it also allows them to estimate the chosen crop's yield rate. It is possible to manually execute these processes. We take the probability values of instances into account here. For new scenarios, we will obtain the outcome. The likelihood of good and evil can be found by the SVM technique. And we can predict whether the chosen crop produces good yields or poor yields. Similarly, the ANN method can measure and find the minimum value for the distance between two values given to the instances. To get the distance between two values, this approach uses Euclidian distance

Datasets

Dataset Collection: We gather data from multiple sources in this process and prepare datasets. And these datasets are used for (descriptive and diagnostic) analytics. Several online repositories of abstracts are open, such as Data.gov.in and indiastat.org. We can use annual crop abstracts for a period of at least ten years. These datasets commonly support anarchic time series conduct. Main and required abstracts (data obtained for at least 10 years) combined-for sugar cane . Original agricultural statistics (specific to sugarcane) and data on rainfall. Moisture, potassium , sodium, phosphorous abundance suitable for sugarcane crop, minimum and best crop temperature, morning and evening RH ethics, Clay parameter (soil depth, PH, edible appropriate acclimatised for crop see table below), acclimatic prediction, sawing date, attacked diseases & sugarcane crop acclimatised pesticides / fertilisers, crop in tones / acres,

As a function of monthly temperature and rainfall, the proposed method should match a regression model that will help us forecast a crop 's yield. Regression would be used by the system because we need to forecast performance is important. Also, the yield will tend to decline in harsh environmental circumstances, which rules out linear regression. Multivariate polynomial regression would then act as the best methodology to suit the relevant model precisely. The next intuitive step, once the model has been created, is to check the consistency of the model i.e. how well the model fits the data given. In essence,this is achieved by calculating residual values. It is essentially possible to describe residuals as leftovers from computed model fit[5].is the residual value (e).

$e = y-\hat{y}$

Then, it determined the remaining number of squares. The error between the data and the fitted model is represented. Its meaning is from 0 to 1 [5]. When applying this model to our dataset, the R-squared value found was calculated to be 0.89.

Results Analysis

An RMSE value of 9.4 was given by multivariate polynomial regression. SVM regression was used in order to increase method accuracy. So, the help vector regression tries to match a hyperplane on the specified dataset. The hyperplane is equipped in such a way that the model is best suited for future forecasting. Also, the vectors defined for hyperplanes are nonlinear for our system. A cost parameter of 4 is used for fine-tuning of the model to control the effect of each human support vector. For evaluating the variance and effect of the support vectors, the gamma (c) value is used. Value c is selected as

$$\gamma = \frac{1}{No\ of\ predictors}$$

The model uses 20 monthly rainfall and temperature predictors. combinations. Consequently, c = 0.05 The epsilon value defines a tolerance margin where errors are not penalised. The larger the epsilon (e), the higher the number of errors the model allows. Multiple e values were checked, and it show the parameter. Value of e is 0.05 the model performs well.

A python application is generated in order to predict the crop yield rate. Three sections are included in this application. First, datasets are managed, second, datasets are tested, and third, datasets are analysed. We can get the datasets of previous years in the management of datasets and they can also be converted into supportive formats In this project, since we use the python library, all datasets are translated to the attribute relationship file format. We may do a single test in the research part. We have considered two machine learning approaches ... We may choose all of the strategies of testing and test data sets, such as choosing a specific crop, a specific position and a specific season, to obtain yield outcomes. We can enter a whole dataset file in the analysis portion.

We were perform the experiment using python and selected the MP Chemical Fertilizer Data Accuracy for classification models is one of the most widely applied indicators. It is the sum of the accurate forecasts divided by the total number of forecasts. For an imbalanced dataset, we can get a high degree of precision that is mostly class-oriented. In an extreme case, each test case could be added to the large class by the classified to obtain a consistency equivalent to that of the most frequent marks in the test set. Precision can therefore be a deceptive measure of performance. The controlled precision is a better measure of generalizability. Where l is the class count, the average accuracy obtained on each class can be calculated.

$$\frac{\sum_i^l (\mathrm{TP}_i+\mathrm{TN}_i)(\mathrm{TP}_i+\mathrm{FP}_i+\mathrm{TN}_i+\mathrm{FN}_i)}{l}$$

Table 1: comparative Analysis different machine learning model

Sample Chemical Fertilizer Datasets	Model	Accuracy (in %)	
		Before artifact removal	After artifact removal
Dataset1	SVM	78.33	88.90
Dataset2	ANN	89.55	90.66
Dataset3	Proposed approach SVM+ ANN	92.12	95.77

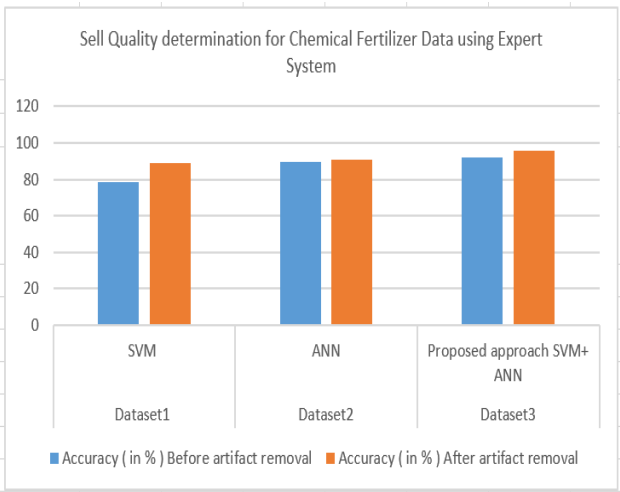


Figure 3: comparative Analysis different machine learning model

6. Conclusion

We proposed hybrid algorithm predicted higher precision wheat yields, and the best generalisation ability of the three approaches was demonstrated by the support vector regression + ANN algorithm. In Madhya Pradesh, the RF model will predict wheat yields reliably in advance (before harvesting dates). We have also observed that the precision of the forecast varies according to agricultural zones and algorithms, and the geographical variation would influence the accuracy of the yield prediction. Such approaches also aid in solving agricultural problems. By testing for various methods, we can also get the precision of yield. We may also boost efficiency by testing the accuracy of various crops. In several agricultural fields, sensor innovations are being applied. This paper aims to bring the crops to their full yield point. It also assists in the selection of the required crop for the chosen land and season. Such methods can address farmers' problems in the field of agriculture. This would continue to boost our country's economic growth.

References

1. Filippi, P., Jones, E. J., Wimalathunge, N. S., Somarathna, P. D. S. N., Pozza, L. E., Ugbaje, S. U., … Bishop, T. F. A. (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precision Agriculture. doi:10.1007/s11119-018-09628-4
2. Shah, A., Dubey, A., Hemnani, V., Gala, D., & Kalbande, D. R. (2018). Smart Farming System: Crop Yield Prediction Using Regression Techniques. Proceedings of International Conference on Wireless Communication, 49–56. doi:10.1007/978-981-10-8339-6_6
3. Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R. L., & Mouazen, A. M. (2016). Wheat yield prediction using machine learning and advanced sensing techniques. Computers and Electronics in Agriculture, 121, 57–65. doi:10.1016/j.compag.2015.11.018
4. Medar, R., Rajpurohit, V. S., & Shweta, S. (2019). Crop Yield Prediction using Machine Learning Techniques. 2019 IEEE 5th International Conference for Convergence in Technology (I2CT). doi:10.1109/i2ct45611.2019.9033611 .

5. Meeradevi, & Salpekar, H. (2019). Design and Implementation of Mobile Application for Crop Yield Prediction using Machine Learning. 2019 Global Conference for Advancement in Technology (GCAT). doi:10.1109/gcat47503.2019.8978315
6. Kumar, R., Singh, M. P., Kumar, P., & Singh, J. P. (2015). Crop Selection Method to maximize crop yield rate using machine learning technique. 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM). doi:10.1109/icstm.2015.7225403.
7. [7]. Vagh, Y., & Xiao, J. (2012). Mining temperature profile data for shire-level crop yield prediction. 2012 International Conference on Machine Learning and Cybernetics. doi:10.1109/icmlc.2012.6358890 .
8. Johnson, M. D., Hsieh, W. W., Cannon, A. J., Davidson, A., & Bédard, F. (2016). Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. Agricultural and forest meteorology, 218, 74-84.
9. Rao, K. R., & Josephine, B. M. (2018, October). Exploring the Impact of Optimal Clusters on Cluster Purity. In 2018 3rd International Conference on Communication and Electronics Systems (ICCES) (pp. 754-757). IEEE.
10. W. Zhuo et al., "Assimilating SAR and Optical Remote Sensing Data into WOFOST Model for Improving Winter Wheat Yield Estimation," 2018 7th International Conference on Agro-geoinformatics (Agro-geoinformatics), Hangzhou, 2018, pp. 1-5.
11. DOI: 10.1109/Agro-Geoinformatics.2018.8476074.
12. H. Aghighi, M. Azadbakht, D. Ashourloo, H. S. Shahrabi, and S. Radiom, "Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, no. 12, pp. 4563-4577, Dec. 2018.
13. DOI: 10.1109/JSTARS.2018.2823361.
14. X. Huang, J. Liu, C. Atzberger and Q. Liu, "Research on the Optimal Thresholds for Crop Start and End of Season Retrieval from Remotely Sensed Time-Series Data Based on Ground Observations," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, 2018, pp. 7727-7730. DOI: 10.1109/IGARSS.2018.8519031.
15. S. Rajawat and A. R. Upadhyay, "Web Personalization Model Using Modified S3VM Algorithm For developing Recommendation Process," 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 2020, pp. 1-6, doi: 10.1109/IDEA49133.2020.9170701.
16. Scholz R.W., Roy A.H., Hellums D.T. (2014) Sustainable Phosphorus Management: A Transdisciplinary Challenge. In: Scholz R., Roy A., Brand F., Hellums D., Ulrich A. (eds) Sustainable Phosphorus Management. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-7250-2_1
17. Xie, Y.W., Yang, J.Y., Du, S.L. et al. A GIS-based fertilizer decision support system for farmers in Northeast China: a case study at Tong-le village. Nutr Cycl Agroecosyst 93, 323–336 (2012). https://doi.org/10.1007/s10705-012-9519-8
18. Liu, N., Li, X. & Waddington, S.R. Soil and fertilizer constraints to wheat and rice production and their alleviation in six intensive cereal-based farming systems of the Indian sub-continent and China. Food Sec. 6, 629–643 (2014). https://doi.org/10.1007/s12571-014-0377-x.
19. Bado, V.B., Djaman, K. & Mel, V.C. Developing fertilizer recommendations for rice in Sub-Saharan Africa, achievements and opportunities. Paddy Water Environ 16, 571–586 (2018). https://doi.org/10.1007/s10333-018-0649-8
1. Wang, Y., Li, C., Li, Y. et al. Agronomic and environmental benefits of nutrient expert on maize and rice in Northeast China. Environ Sci Pollut Res 27, 28053–28065 (2020). https://doi.org/10.1007/s11356-020-09153-w