ANALYSIS OF MOISTURE LEVEL IDENTIFICATION OF RED SOIL IN TERMS OF WEATHER CONDITION USING DATA MINING ALGORITHMS.

Mr.P.Gowthamraj[1], Mr.M.Rajkumar[2], Mr.T.Bharathkumar[3]

III BCA, Kaamadhenu Arts&Science College, Sathyamangalam[1]

Assistant Professor,Department of CA&IT, Kaamadhenu Arts and ScienceCollege,Sathyamangalam[2]

III BCA, Kaamadhenu Arts and Science college ,Sathyamangalam[3]

Abstract : Soil is an important key issue of agriculture. The target of the work is to predict soil sort using data processing classification techniques. Soil sort is expected victimization data processing classification techniques like JRip, 148 and Naive mathematician. These classifier algorithms are applied to extract the data from soil data and 2 forms of soil are thought-about like Red and Black. During this paper, data processing and agricultural Data Mining are summarized. The JRip model will manufacture a lot of reliable results of this knowledge and therefore the alphabetic character Statistics within the forecast was accrued. For finding the problems in huge knowledge, economical strategies may be created that utilize data processing to boost the exactitude of classification of big soil knowledge sets.

1. INTRODUCTION:

Data Mining (DM) is famous in the land subdivision, wasteland management and pest and pest control. In 1 tested a variety of integration techniques in DM and applied on the website
of soil science to predict meaningful relationships and provide for the rules of association for different types of soil in agriculture. Similarly, agricultural forecasts, diseases the detection and efficacy of pesticides are analyzed using various data mining techniques before2. In3 analyzes the J48 split algorithm with high accuracy predicting soil fertility rate. In4 investigates usage of the various DM methods of obtaining information on in the agricultural sector and presented various exhibitions to obtain information in the manner of the Organization Rules, Consolidation, Separation and Relationships. By5 predicted levels of soil fertility using differentiation strategies were Naïve Bayes, J48 and K-Nearest Neighbor algorithms. In6 used data mining techniques Accepted to measure the analysis of crop yields. Multiple Row Route (MLR) method used to find line relationships between dependent and independent variables. K-Means the combination method was also used to create four sets considers Rain as the primary parameter. In7 analyzed I features of landslides in the Shimen lake water area in northern Taiwan. Decision tree, Bayesian Network and non-network data mining methods were used. Improvement is supported The Bayesian Network approach has been considered better than that non-linear. In8 analyzed the apparent value of land reproduction and crop management features to predict maize yield and determining crop yields as well the gap between farmers. **Separation** and descent tree analysis was used to predict the outcome. In9 is investigating two broad methods of calculating production the associated yield gap and the nutrient associated with soil fertility balance. The system allows information from micro- we measure to a higher level and determine world quality. In10 predicted soil characteristics and analyzed soil data using separation strategies. Soil properties such as pH value, Electrical Conductivity (EC), Potassium, iron, copper, etc. they were classified using classification algorithms such as Naïve Bayes, J48 and JRip. Among the algorithms, J48 was present is considered a simple separator and produced a better result.

2. DATA MINING IN AGRICULTURE

Data mining is essential to uncover agriculture-related knowledge such as soil fertility, yield prediction, and soil erosion. Soil prediction helps in soil remediation and crop management. Classification algorithms involve finding rules that separate data into separate groups. Such a classification process generates a set of classification rules that can be used to classify future data11.The next section contains an explanation of classification algorithms such as the Naive Bayes classifier, the J48 decision tree classifier, and the JRip classifier.

2.1 Navie Bayes

A Naive Thomas Bayes category fier is one among one amongst one in every of the classifiers in a family of straightforward probabilistic classification techniques in machine learning. It's supported the Bayes theorem with independence features. Every class labels are calculable through chance of given instance. It wants solely tiny quantity of coaching information to predict class label necessary for classification.

2.2 J48

The J48 is one in every of the classification-decision tree formula and it slightly changed from C4.5 in Weak. It will choose the check as best info gains. This algorithm was planned by Ross Quinlan. C4.5 is additionally cited as a applied math classifier. J48 predicts variable from accessible data. It builds tree supported attributes values of coaching data. This classifies data with the assistance of feature of information instances that said to possess information gain. The importance of error tolerance is developed victimization pruning concept13,14.
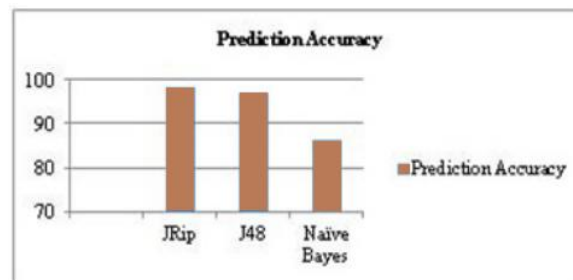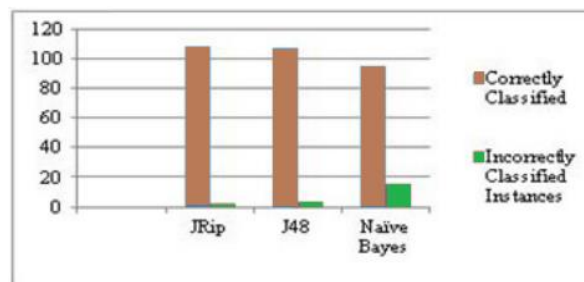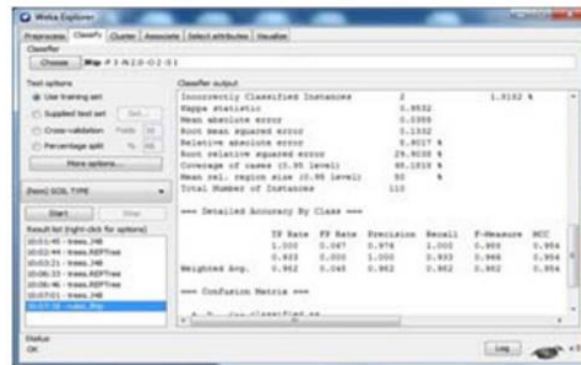
2.3 IREP

IREP optimized model is Repeated Incremental Pruning to Produce ErrorReduction (RIPPER), which turned into proposed through William W. Cohen. This set of rules is a propositional guiding principle learner. J-Rip classifier is one of the choice tree pruning fashions primarily based totally on association policies. It is a powerful approach to lessen blunders pruning. In this set of rules, the education statistics is break up into units and with the assist of pruning operators the mistake is decreased on each the units. Finally policies are shaped from units which include Growing set and Pruning set.

3. OVERALL IDEA OF THE PAPER

In this work, we collect the agricultural soil dataset from the Soil Testing Laboratory, Virudhunagar District. We took 110 data containing the attributes such as such as village name, soil type or color, soil texture, PH, EC (electrical conductivity), condition of lime, phosphorus. This system predicted the type of red and black soil based on the values of PH and EC. The PH of bare black soil is higher than 7.7 and that of red soil is lower than 7. We use three classification algorithms like JRip, J48, and Naive Bayes to predict the type of red and black soil. By applying three classification algorithms, JRip considers the entire attribute. However, the J48 classifier only considers the value of PH and EC. The tree is built based on the previous two attributes. The JRip classifier generates the rules efficiently and performs well for this soil dataset. If you compare these three JRip algorithms resulted in high accuracy. Here, the entire dataset is considered a training set. Based on the training data set, we conclude that the weighted average is the true positive rate of the JRip classifiers 0.982. In the case of J48 and Naive Bayes, the TP rate is 0.97 and 0.86 indicates the low level. So the JRip classifier automatically classified the data set in a higher sense. Soil properties differed between sites with red textured soil sand black textured floors. Since below 7.0 is acidic soil and above 7.0 is alkaline soil. The spectral analysis was sensitive enough to detect differences in soil fertility between different soil types. The soil data set with attributes such as soil type, pH value, etc. are given in Figure 1. This data set organized in an Excel sheet with saves as file type is CSV extension.

4.   FUTURE SCOPE AND CONCLUSION

This article projects the comparative analysis of three algorithms such as Naive Bayes, JRip and J48. The JRip sorting algorithm performs better on this dataset and sorts correctly at the maximum number of instances compared to the other two. JRip can be recommended to predict soil types.

5.   REFERENCE

1. Geetha MCS. Implementation of association rule mining for different soil types in agriculture. International Journal of Advanced Research in Computer and Communication Engineering. 2015 Apr; 4(4):520–2.
2. Solanki J, Mulge Y. Different techniques used in data mining in agriculture. International Journal of Advanced Research in Computer Science and Software Engineering.2015 May; 5(5):1223–7.
3. Bhuyar V. Comparative analysis of classification techniques on soil data to predict fertility rate for Aurangabad District.International Journal of Emerging Trends and Technology in Computer Science. 2014 Mar-Apr; 3(2):200–3.
4. Fathima NG, Geetha R. Agriculture crop pattern using data mining techniques. International Journal of Advanced Research in Computer Science and Software Engineering.2014 May; 4(5):781–6.
5. Suman, Naib BB. Soil classification and fertilizer recommendation using WEKA.International Journal of Computer Science and Management Studies. 2013 Jul; 13(5):142–6.
6. Ramesh D, Vardhan VB. Data mining techniques and applications to agricultural yield data. International Journal of Advanced Research in Computer and Communication Engineering. 2013 Sep; 2(9):3477–80.
7. Tsai F, Lai JS, Chen WW, Lin TH. Analysis of topographicand vegetative factors with data mining for landslide verification.Ecological Engineering.2013 Dec; 61:669–77.

8. Tittonell P, Shephered KD, Vanlauwe B, Giller KE. Unravel-ling the effects of soil and crop Management on maize productivity in small holder agricultural systems of Western Kenya - An application of classification and regression tree analysis. Agriculture, Ecosystems and Environment. 2008 Jan; 123(1-3):137–50.
9. Bindraban PS, Stroorvofel JJ, Jansen DM, Vlaming J, Groot JJR. Land quality indicators for suitable land management:Proposed methods for yield gap and soil nutrient balance.Agriculture, Ecosystems and Environment. 2000; 81:103–
10. Gholap J, Lngole A, Gohil J, Shailesh, Attar V. Soil dataanalysis using classification techniques and soil attribute prediction. 2012 Jun; 9(3):1–4.
11. Anuradha C, Velmurugan T. A comparative analysis on the evaluation of classification algorithms in the prediction of student performance. Indian Journal of Science and Technology. 2015 Jul; 8(15):1–12.
12. Narain B. Study for Data Mining techniques in classification of agricultural land soils. Journal of Advanced Research in Computer Engineering. 2011 Jan-Jun; 5(1):35–7.
13. Venkatesan E, Velmurugan T. Performance analysis of decision tree algorithms for breast cancer classification. Indian Journal of Science and Technology. 2015 Nov; 8(29):1–8.
14. Chandrakar PK, Kumar S, Mukherjee D, Applying classification techniques in Data Mining in agricultural land soil. International Journal of Computer Engineering. 2011 JulDec; (2):89–95.