Abstract

Agriculture is the backbone of the Ethiopian economy and it contributes the highest GDP of the country. Among this, crop production takes the highest level of income for most smallholder farmers in all regions of Ethiopia.

The objective of this research is to build a model that can predict crops productivity and implement a decision support system. In order to conduct this research, a hybrid Knowledge Discovery Process model was adopted. For the purpose of this research, the datasets were taken from Central Statistical Agency of Ethiopia database, and the researcher used a total of 25,000 instances for training and building a model. Hence, for building a model and implementing decision support system for predicting crop productivity, WEKA data mining tool and java NetBeansIDE was used respectively. To achieve the objective of these research different experiments were conducted using J48, HoeffdingTree decision tree and PART rule based classifiers. In addition, the predictive performances of the classifiers are evaluated and
Application of Data Mining Tools for Identifying Determinant Factors for Crop Productivity

compared using accuracy rate, confusion matrix and ROC curve. Based on this, out of the three classifiers PART rule based classifier performs best accuracy and ROC rate which is 95.44 % and 0.992 respectively. As a result PART rule based classifier were selected for implementing the model to predict crop productivity. In this thesis, the experimental result shows that, the main determinant factors for crop productivity are main season (season type), use of extension program, fertilizer used and fertilizer type. Therefore, the outcome of this research is essential to make data mining based decisions for policy makers and for experts in the area of crop agriculture to give an attention on the factors affecting crop productivity and to take corrective measures.

References

1. K. M, Data Mining: Concepts, Models, Methods, and Algorithms. John Wiley & Sons Inc., 2003.
2. H. J and K. M, Data mining, concepts and techniques, 2nd ed. Morgan Kaufmann Pub, 2006.
3. Z. Yan-li and Z. Jia, "Research on Data Preprocessing In Credit Card Consuming Behavior Mining", Energy Procedia, vol. 17, pp. 638-643, 2012.
4. L. Xiang-wei and Q. Yian-fang, "A Data Preprocessing Algorithm for Classification Model Based On Rough Sets", Physics Procedia, vol. 25, pp. 2025-2029, 2012.
5. A. Seyoum, P. Dorosh and S. Asrat, "Crop Production in Ethiopia: Regional Patterns and Trends", Ethiopian development research institute, 2011. [Online]. Available: http://reliefweb.int/sites/reliefweb.int/files/resources/essprn11.pdf. [Accessed: 09- Jan 2016].
6. D. Iadeni , Data mining and decision support. Boston: Kluwer Academic publishers, 2003.
7. D. Iadeni , Data mining and decision support. Boston: Kluwer Academic Publishers, 2003.
8. K. Cios, L. Kurgan, W. Pedrycz and R. Swiniarski, Data mining. New York, NY: Springer, 2007.
9. V. S, "Crop productivity mapping based on decision tree and Bayesian classification", 62 Unpublished, 2007.
10. S. Sawaitul, P. Wagh and D. Chatur, "Classification and Prediction of Future Weather by using Back Propagation Algorithm- An Approach", International Journal of Emerging Technology and Advanced Engineering, vol. 2, no. 1, 2012
11. D. . and B. Vardhan, "ANALYSIS OF CROP YIELD PREDICTION USING DATA MINING TECHNIQUES", International Journal of Research in Engineering and Technology, vol. 04, no. 01, pp. 470-473, 2015.
12. T. Bekele, "building a predictive model for annual cereal crops production using data mining techniques", 2014
13. Z. Diriba and B. Borena, "Application of Data Mining Techniques for Crop Productivity Prediction", HiLCoE Journal of Computer Science and Technology, vol. 1, no. 2.

Index Terms

Computer Science

Data Mining
Keywords

Data mining, predictive model, decision support system, crop production, Ethiopia