Authors:

Subi Jain, Gagan Vishwakarma, Yogendra Kumar Jain

Abstract

Decision-making process is supported by Machine learning-based classification techniques in many areas of health care. Classification performance of decision system can be improved using the attribute reduction mainly in the situation of high data dimensionality dilemma .This paper proposes, Random forest Classifier (RFC) approach which is based on the Variable Precision Rough Set (VPRS) theory. The first phase of proposed approach focus at attribute reduction of available dataset using VPRS .Directing from dimensionality reduction to predictive model construction, and in next phase, the obtained abridged dataset is provided as the input of RFC to build a more accurate classification model. The performance is evaluated in terms of classification accuracy and time complexity. The experimental results show that the enhanced RFC has higher accuracy and correctly classified instances as compared with the existing algorithms.

References

1. Pawlak, "Rough sets", International Journal of Computing. Information Sciences, vol.11, pp.341-345, 1982.

2. QiangShen and Richard Jensen, "Rough Sets, their Extensions and Applications," International Journal of Automation and Computing, 04(1), pp. 100-106, 2007.

3. J .R .Quinlan." Induction of Decision Trees." Machine Learning. 1, 81-106, 1 (Mar. 1986).

4. T. Ho, "Random decision forest, in: 3rd International Conf. on Document Analysis and Recognition", pp. 278–282, 1995.

5. Y. Amit, D. Geman , "Shape quantization and recognition with randomized trees", Neural Computing. 9, 1997.

6. L. Breiman, "Random forests", Mach. Learn. 45, pp 5–32, 2001.

7. K. Polat, S. Gunes, "A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems", Expert Syst. App. .: Int. J. 36 , 2009.

8. V.F. Rodriguez-Galiano, B. Ghimire , J. Rogan , M. Chica-Olmo , J.P. Rigol-Sanchez," An assessment of the effectiveness of a random forest classifier for land-cover classification", Elsevier J. of ISPRS Journal of Photogrammetry and Remote Sensing 67, pp 93–104,2012.

9. Alexander Hapfelmeier , "Analysis of Missing Data with Random Forests",2012.

10. Ahmad TaherAzar, Hanaa Ismail Elshazly, Aboul Ella Hassanien,Abeer Mohamed Elkorany",A random forest classifier for lymph diseases",Elsevier J. of computer methods and programs in biomedicine 113,pp 465–473,2014.

11. Mojtaba Seyedhosseini, Tolga Tasdizen, "Disjunctive normal random forests", Pattern Recognition, vol 48, pp 976–983, 2015.

12. UCI. Machine Learning Repository. http://archive.ics.uci.edu/ml/index.html

13. Faraz Akram, Seung Moo Han, Tae-Seong Kim, "An efficient word typing P300-BCI system using a modified T9interface and random forest classifier", Computers in Biology and Medicine, vol. 56 ,pp 30–36,2014.

14. CHEN Jiajun, HUANG Yuanyuan, "Decision Tree Construction Algorithm for Incomplete Information System", IEEE fourth International Conference on Computational and Information Sciences, pp 404-407, 2012.

15. In-Kyoo Park a , Gyoo-Seok Choi ," A variable-precision information-entropy rough set approach for job searching" , Elsevier J. of Information Systems 48 ,279–288, 2015.

Index Terms

Computer Science   Artificial Intelligence

Keywords

Decision tree classification, Attribute reduction, Variable Precision Rough Set (VPRS), Random forest classifier (RFC).