Abstract

Machine learning is a concerned with the design and development of algorithms. Machine learning is a programming approach to computers to achieve optimization. Classification is the prediction approach in data mining techniques. Decision tree algorithm is the most common classifier to build tree because of it is easier to implement and understand. Attribute selection is a concept by which be select more significant attributes in the given datasets. These proposed a novel hybrid approach combination of VPRS with Boundary Region and Random Forest algorithm called VPRS Boundary Region based Random Forest Classifier (VPRSBRRF Classifier) which is used to deal with uncertainties, vagueness and ambiguity associated with datasets. In this approach, select significant attributes based on variable precision rough set theory with boundary region as an input to Random Forest classifier for constructing the decision tree which is more efficient and scalable approach for classification of various datasets.

2. Study on discretization in rough set based on genetic algorithm cai-yun chen, zhi-guo li, sheng-yong qiao, shuo-pin wen Center for Combinatorics, LPMC, Nankai University, Tianjin, 300071, P. R. China E-MAIL: sbickle@eyou.com


5. Rough Sets, their Extensions and Applications Qiang Shen, Richard Jensen. 04(1), January 2007, 100-106 DOI: 10.1007/s10453-004-5872-7


8. KDD Cup 99 Dataset, 1999.


13. A Comparative Study of Discretization Methods for Naive-Bayes Classifiers Ying Yang & Geoffrey I. Webb2 1 School of Computing and Mathematics Deakin University, VIC 3125, Australia yyang@deakin.edu.au 2 School of Computer Science and Software Engineering Monash University, VIC 3800, Australia Geoff.Webb@mail.csse.monash.edu.au


Index Terms

Computer Science

Algorithms
Keywords

Discretization, Variable Precision Rough Sets, Boundary Region, Random Forest