Abstract

A non-parametric, very simple to use, effective instance-based learning algorithm called K-Nearest Neighbor (KNN), is most widely used to classify the objects in data mining. KNN has some shortcomings which affect its classification performance like the equal impact of all attributes, curse of dimensionality, the value of ‘k’ parameter and simple voting. A variety of techniques are developed in literature to get better performance. This paper presents an improved algorithm called dual weighted KNN that is a combination of attribute weighted and instance weighted techniques. To verify the performance of proposed algorithm it is conducted on fourteen different datasets taken from UCI in ‘R’ data mining tool. The results show that the proposed algorithm significantly outperforms than traditional KNN, Attribute weighted KNN and Instance weighted KNN.

References

1. J. Han and M. Kamber, Data mining: concepts and techniques, 2nd ed. Amsterdam;
Evaluating the Performance of Dual Weighted K-Nearest Neighbor Classifier


Index Terms

Computer Science | Information Sciences
Keywords

KNN, Attribute weighted KNN, Instance weighted KNN, and Distance weighted KNN.