Abstract

Because of the massive increase in the size of the data it becomes troublesome to perform effective analysis using the current traditional techniques. Big data put forward a lot of challenges due to its several characteristics like volume, velocity, variety, variability, value and complexity. Today there is not only a necessity for efficient data mining techniques to process large volume of data but in addition a need for a means to meet the computational requirements to process such huge volume of data. The objective of this paper is to classify big data using Fuzzy K-Nearest Neighbor classifier, and to provide a comparative study between the results of the proposed systems and the method reviewed in the literature. In this paper we implemented the Fuzzy K-Nearest Neighbor method using the MapReduce paradigm to process on big data. Results on different data sets show that the proposed Fuzzy K-Nearest Neighbor method outperforms a better performance than the method reviewed in the literature.

References


10. Zhiqiang Liu; Hongyan Li ; Gaoshan Miao.MapReduce-based Backpropagation Neural Network over large scale mobile data

11. Changlong Li1, Xuehai Zhou1, Kun Lu1. Implementation of Artificial Neural Networks in MapReduce Optimization.


Index Terms

Computer Science Fuzzy Systems
Keywords

Big data; Classification; Fuzzy k-nearest neighbor; Fuzzy logic; Hadoop; MapReduce