A Progressive Review on Feature Subset Selection for High Dimensional Data Analytics

Archana Sumant

MET's Institute of Engineering Nashik , India,

archana.vaidya@ges-coengg.org

Dr. D. V. Patil

2GES's R. H. Sapat College of Engineering Management Studies and Research Nashik, India,

dipak.patil@ges-coengg.org

Abstract: The emanation of the big data induces severe problems and challenges besides the apparent benefits. Industry demands and cutthroat pressure nearly every business has a huge requirement for processing of data in short time and with more predictive accuracy. Accordingly, from big data perspective mining of valuable information from enormous amount of data competently and accurately is a big problem. The big data has characteristics such as high dimensionality, versatility, variability, veracity and noise. Massive data regularly hold properties found in different information factors in hundreds or thousands of levels. Most of the variables from this data may contain very less information. Another problem is to apply suitable strategies to get good classification performance for a high dimensional dataset .In this paper we have done a progressive review on variable subset selection to deal with high dimensional data and proposed a new approach with hybrid model and extreme learning machine classifier to increase the predictive accuracy of classifier and parallel approach used reduces the time complexity of computation.