TECHNOLOGIES

abtechchennai@gmail.com
www.abtechnologies.in
9840511458

SPLITTING LARGE MEDICAL DATA SETS BASED ON NORMAL DISTRIBUTION IN CLOUD ENVIRONMENT

ABSTRACT

The surge of medical and e-commerce applications has generated tremendous amount of data, which brings people to a so-called “Big Data” era. Different from traditional large data sets, the term “Big Data” not only means the large size of data volume but also indicates the high velocity of data generation. However, current data mining and analytical techniques are facing the challenge of dealing with large volume data in a short period of time. This project propose the efficiency of utilizing the Normal Distribution (ND) method for splitting and processing large volume medical data in cloud environment, which can provide representative information in the split data sets. The ND-based new model consists of two stages. The first stage adopts the ND method for large data sets splitting and processing, which can reduce the volume of data sets. The second stage implements the ND-based model in a cloud computing infrastructure for allocating the split data sets. The experimental results show substantial efficiency gains of the proposed method over the conventional methods without splitting data into small partitions. The ND-based method can generate representative data sets, which can offer efficient solution for large data processing. The split data sets can be processed in parallel in Cloud computing environment.