ملخص البـــــحث باللغة الانجليزية :

Many scholars and practitioners in industrial system engineering and cybernetics are interested in "big data." Big data differs from regular data. Big data, for example, is often unstructured and has to be analyzed in real-time. Big data analytics has received considerable attention from business and academia because of its tremendous cost-cutting and decision-making advantages. For example, it can provide essential information for a wide range of businesses. In this paper, we developed a decision tree (DT) algorithm, specifically the classification and regression trees, to obtain better results. We designed a modified version of the DT algorithm. The tree aims to achieve optimal depth by self-tuning running parameters and improving the performance. Furthermore, we verified the efficiency and efficacy of the modified DT using two datasets: airport and fire datasets. The airport and fire datasets contain 500000 and 600000 instances, respectively. We also compared the modified and standard DT algorithms. The results showed that the modified DT algorithm performs better than the standard DT algorithm. This comparison was conducted on multi-node on the Apache Spark tool using Amazon web services. Moreover, we achieved an accuracy of 6.85% and 8.85% for the fire and airport datasets, respectively.

البحث غير مشتق من رسالة علمية

يقع البحث ضمن مجالات البحث بالقسم العلمي

عميد الكلية

أ.د/ محمد حلمي عبد العزيز خفاجي