Redundancy Avoidance for Big Data in Data Centers: A Conventional Neural Network Approach

Abstract—As the innovative data collection technologies are applying to every aspect of our society, the data volume is skyrocketing. Such phenomenon poses tremendous challenges to data centers with respect to enabling storage. In this paper, a hybrid-stream big data analytics model is proposed to perform multimedia big data analysis. This model contains four procedures, i.e., data pre-processing, data classification, data recognition and data load reduction. Specifically, an innovative multi-dimensional Convolution Neural Network (CNN) is proposed to assess the importance of each video frame. Thus, those unimportant frames can be dropped by a reliable decision-making algorithm. In order to ensure video quality, minimal correlation and minimal redundancy (MCMR) are combined to optimize the decision-making algorithm. Simulation results show that the amount of processed video is significantly reduced, and the quality of video is preserved due to the addition of MCMR. The simulation also proves that the proposed model performs steadily and is robust enough to scale up to accommodate the big data crush in data centers.

CONCLUSIONS

In this paper, a hybrid-stream big data analytics model has been proposed to enhance the classification precision and relieve the data centers’ network and storage overload. The model can improve the speed to deal with the videos and recognizing, deciding the important frames and whether to drop the unimportant ones in every video. Compared to conventional methods like deep learning to address image analysis problems, this paper has improved the method to deal with video analysis. Besides, this network and storage overload problem of video is considered as an optimization problem, which can show a practical algorithm over a largescale of real-time data from numerous nodes.
SYSTEM REQUIREMENTS:

HARDWARE REQUIREMENTS:

- System: Pentium IV 2.4 GHz.
- Hard Disk: 40 GB.
- Floppy Drive: 1.44 Mb.
- Monitor: 15 VGA Colour.
- Mouse: Logitech.
- Ram: 512 Mb

SOFTWARE REQUIREMENTS:

- Operating system: Windows 7/UBUNTU.
- Coding Language: Java 1.7, Hadoop 0.8.1
- IDE: Eclipse
- Database: MYSQL

REFERENCES

