Performance Evaluating System Based on MapReduce in Context of Educational Big Data

Chitresh Verma, Amity University, Lucknow, India
Rajiv Pandey, Amity University, Lucknow, India
Devesh Katiyar, D.S.M.N.R.U., Lucknow, India

ABSTRACT

The Big Data subsist in every characteristic of our daily life. The educational Big Data is one of these aspects of Big Data which is linked to student life. This article provides the comprehensive understanding of the implementation of the grade analysis system using educational Big Data. The grade analysis can be used for helping the students in many ways like selecting an elective subject, determining the toughness of elective subjects. Selection of good subject by the student will improve their career opportunities and placement probability. Further, this article builds a reference performance evaluating system for any future performance system with other aspects like employee performance evolution in any human capital management system (HCM).

KEYWORDS

Big Data Computation, Big Data, Grade Analysis, Hadoop, HDFS, JAVA, MapReduce, Student Performance

1. INTRODUCTION

Big Data had altered the way we had contemplation of the data and when combined with Omni presences with Big Data along with its influence on our daily life. One of these influences of Big Data is education. We can refer to Big Data in the educational domain as Educational Big Data. Educational Big Data can be processed by a framework for Big Data called Hadoop. Hadoop has been developed as part of Apache Hadoop project (The Apache Software Foundation, 2016) as open-source software. It consists of many components and sub components. The main components of Hadoop are Hadoop MapReduce, HDFS, and Hadoop YARN. (The Apache Software Foundation, 2016)

In brief, Hadoop MapReduce is programming model for Hadoop. MapReduce is a two-stage process of programming model. These two stages are Mapper stage and Reducer stage. In Mapper stage, all the raw data is transformed to key-value pairs. The key-value pair is a combination of two data, such that one of data can be used to identify the other data. In Reducer stage, data are computed as per the logic of reducer. Furthermore, the Hadoop makes use of the JAVA programming language and it is one of the most popular languages in the world. JAVA provides various predefined libraries for computation. It is also important to point out that Hadoop developers make use of JAVA language for writing MapReduce jobs.

A system involving analysis of the marks of the students, after which certain conclusion about the data can be attained can be developed using the Hadoop and Educational Big Data. This system can be termed as Grade Analysis. In other words, Grade Analysis can be understood as the process

DOI: 10.4018/IJOCI.2018010101

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
involving storage and analysis of the marks or the grade of the student by a computational framework to do any specific task.

This paper tries to explore means of grade analytics implementation on Hadoop to help the student in selecting elective subject. To enable better understanding the paper has been divided into 10 sections. First section provides a short introduction to the background related to this paper. Second section discusses about the related work in field of grade analysis. The third section provides a brief introduction to Big Data similarly fourth section gives a brief introduction to Hadoop framework. In the fifth section, Grade Analysis is described with its proposed architecture. Storage of Educational Big Data on HFDS is explained in the fifth section. Algorithm based on MapReduce and its implementation is provided in seventh and eighth section respectively. The ninth section provides the analysis associated to result which is obtained after running proposed implementation in actual circumstances. Finally, the tenth section makes the conclusion of the paper and also provides future scope of this paper.

2. RELATED WORK

Many grading systems and studies had been developed over the last decades to help the student in improving their performance. One such system is the literature based learning and performance system which measured along its analysis was developed (Muñiz-Swisegood, 1994). It involved the testing the learning capabilities of students consisting of bilingual and monolingual. The experiment mainly focused on learning of students speaking Spanish and English language. The findings were published in Bilingual Research Journal, 1994 and involved results of the examination of metacognitive instructions learning.

One another grade analysis system involved admission test marks and grade point average (GPA) and it tried to predict performance of the student in Pharmacy College to help the pharmacy colleges and license examiner. (Kuncel, Credé, Thomas et al., 2005) The experiment used the Hunter and Schmidt psychometric meta-analytic method. It had applied the method of comparison of admission test marks and GPA.

Researchers in University of Murcia, Spain had demonstrated applications of the Big Data analytics in implementation of smart cities and their campuses of any university in their paper titled “Applicability of Big Data Techniques to Smart Cities Deployments”. (Victori, 2016) They had explored the potential applications like smart energy management, smart building and smart traffic management systems using IoT and Big Data analysis with multiple scenarios in the university campus.

3. EDUCATIONAL BIG DATA: DATA COLLECTION AND PRE-PROCESSING

The Big Data is the large sets of data which thought to be beyond the capabilities of traditional system to process. (Kuncel, Credé, Thomas et al., 2005) The Big Data is defined on various Vs in their respective ways by the researcher. The Educational Big Data is the dataset which related to student which may be used examine any parameter of his education. For example, the data related to marks of students in the final semester belonging to collections of Indian universities.

Education systems have transformed with the introduction of technology in the learning system. The students are generating a lot of data during their academic session. These are mostly related to their education. This data may be about discussion on any topic amongst the students inside discussion forum or it may involve health related data of student available on college medical care.

Enlargement of the scope in data has transformed the landscape for analysis. There seems to be requirement of using a Big Data Analysis system like Hadoop. The simple operation like count on traditional system can be performed, but combined them with real-time system with various aspects of other data like alignment and create relationship remains a challenge. Educational Big Data could be stored and processed and in following section we will discuss the simple count operation of student grade which could be evolved into multiple attributed data analytics.
Cognitive Bare Bones Particle Swarm Optimisation with Jumps
www.igi-global.com/article/cognitive-bare-bones-particle-swarm-optimisation-with-jumps/144240?camid=4v1a

Cloud Computing: A Practical Overview Between Year 2009 and Year 2015
www.igi-global.com/article/cloud-computing/135981?camid=4v1a