Bing Lin

Data Analytics

11/10/21

Big Data Analytics

The concept of big data analysis has existed for many years. Now, most organizations understand that if they capture all the data flowing into their business, they can apply analytics and derive great value from it. But even in the 1950s, before the term "big data" was spoken decades ago, companies had been using basic analysis (essentially manual inspection of numbers in spreadsheets) to discover insights and trends.

　　 However, the new benefits of big data analysis are speed and efficiency. Years ago, companies could collect information, run an analysis, and dig out information that can be used for future decisions. Today, companies can identify insights for immediate decision-making. The ability to work faster and remain agile provides organizations with unprecedented competitive advantages.

NoSQL is one of the cores of big data analytic. To process data analytics, first, you need to be able to save big data. In the traditional system, HDFS allows you to do that, but after you save this data, you need to think about how to process them. Even though HDFS can fully organize data in different machines but since the data is too big, if you only allow one machine to read all the data, it might take up days even weeks. To improve that, people introduced MapReduce/Tez/Spark functions, but the programmer found that the MapReduce program was troublesome to write. They hope to simplify this process, this is when Pig and Hive take place. With Hive, people found that SQL has a huge advantage over Java. One is that it is too easy to write. The word frequency just now has only one or two lines to describe in SQL, and it takes about dozens or hundreds of lines to write in MapReduce. But this is only the basics of a data warehouse, what if I want the data to process faster? Due to the inherent limitations of traditional relational databases, such as peak performance, scalability, fault tolerance, and poor scalability, it is difficult to meet the flexible management requirements of massive data. For this reason, NoSQL, a new model for massive data management, appeared. Key-value in NoSQL fastens the time it takes to find the data that bonds with a certain key, it may be only a few tenths of a second. This allows some specialized operations of big data companies to be greatly optimized.

An example of big data analytics uses include customer acquisition and retention. Consumer data are collected and exanimated to improve one company’s customer service to increase consumer satisfaction. For example, Amazon and Netflix have personalized engines that provide improved customer experiences to earn customer loyalty.

In the 21st century, when Internet thinking continues to collide, big data analysis will surely become the next important technological innovation. We can focus on big data and understand what big data analysis is and the benefits it can bring us. We know that big data analysis is based on opportunities and challenges, and at the same time, it will change all aspects of our lives in the future and make our lives better.

Works Cited

Chai, W., Labbe, M., & Stedman, C. (2021, February 1). *What is Big Data Analytics and why is it important?* SearchBusinessAnalytics. Retrieved November 10, 2021, from https://searchbusinessanalytics.techtarget.com/definition/big-data-analytics.

Dix, J. (2014, March 11). *What's better for your big data application, SQL or NoSQL?* Network World. Retrieved November 10, 2021, from https://www.networkworld.com/article/2226514/what-s-better-for-your-big-data-application--sql-or-nosql-.html.