Chapter 86
Big Data Analytics Demystified

Pethuru Raj
IBM India Pvt Ltd, India

ABSTRACT

This chapter is mainly crafted in order to give a business-centric view of big data analytics. The readers can find the major application domains / use cases of big data analytics and the compelling needs and reasons for wholeheartedly embracing this new paradigm. The emerging use cases include the use of real-time data such as the sensor data to detect any abnormalities in plant and machinery and batch processing of sensor data collected over a period to conduct failure analysis of plant and machinery. The author describes the short-term as well as the long-term benefits and find and nullify all kinds of doubts and misgivings on this new idea, which has been pervading and penetrating into every tangible domain. The ultimate goal is to demystify this cutting-edge technology so that its acceptance and adoption levels go up significantly in the days to unfold.

INTRODUCTION

Today, besides data getting originated in multiple formats and types, data sources, speeds, and sizes are growing expediently and exponentially. The device ecosystem is expanding fast thereby resulting in a growing array of fixed, portable, wireless, wearable, nomadic, and mobile devices, instruments, machines, consumer electronics, kitchen wares, household utensils, equipment, etc. Further on, there are trendy and handy, implantable, macro and nano-scale, disposable, disappearing, and diminutive sensors, actuators, chips and cards, tags, speckles, labels, stickers, smart dust, and dots being manufactured in large quantities and deployed in different environments for gathering environment intelligence in real time. Elegant, slim and sleek personal gadgets and gizmos are really appealing to people today. Self, situation and surroundings-awareness are being projected as the next-generation feature for any casually found cheap things. With the set of empowerment tasks such as digitalization, service-enablement and extreme connectivity, the future hardware and software systems are bound to exhibit real-time and real-world intelligence in their operations, outputs and offerings. Knowledge extraction, engineering and exposition will become a common affair.

There are deeper and extreme connectivity methods flourishing these days. Integration and orchestration techniques, platforms, and products have matured significantly. The result is that Information and Communication Technology (ICT) infrastructures, platforms, applications, services, social sites, and databases at the cyber level are increasingly interconnected with devices, digitalized objects (smart and sentient materials) and people at the ground level via a variety of networks and middleware solutions. There is a strategic, seamless and spontaneous convergence between the virtual and physical worlds. All these clearly insist the point that data creation / generation, capture, transmission, storage and leverage needs have been growing ceaselessly. This positive and progressive trend is indicated and conveying a lot of key things to be seriously contemplated by worldwide business and IT executives, engineers and experts. New techniques, tips, and tools need to be unearthed in order to simplify and streamline the knowledge discovery process out of data heaps. The scope is bound to enlarge and there will be a number of fresh possibilities and opportunities for business houses. Solution architects, researchers and scholars need to be cognizant of the niceties, ingenuities, and nitty-gritty of the impending tasks of transitioning from data to information and then to knowledge. That is, the increasing data volume, variety, and velocity have to be smartly harnessed and handled through a host of viable and valuable mechanisms in order to extract and sustain the business value.

THE UNWRAPPING OF BIG DATA COMPUTING

We have discussed about the fundamental and fulsome changes happening in the IT and business domains. Service-enablement of applications, platforms, infrastructures and even everyday devices besides the varying yet versatile connectivity methods has laid down strong and simulating foundations for man as well as machine-generated data. The tremendous rise in data collection along with all the complications has instinctively captivated both business and IT leaders to act accordingly to take care of this huge impending and data-driven opportunity for any growing corporates. This is the beginning of the much-discussed and discoursed big data computing discipline. This paradigm is getting formalized with the deeper and decisive collaboration amongst product vendors, service organizations, independent software vendors, system integrators, innovators, and research organizations. Having understood the strategic significance, all the different and distributed stakeholders have come together in complete unison in creating and sustaining simplifying and streamlining techniques, platforms and infrastructures, integrated processes, best practices, design patterns, and key metrics to make this new discipline pervasive and persuasive. Today the acceptance and activation levels of big data computing are consistently on the climb. However it is bound to raise a number of critical challenges but at the same time, it is to be highly impactful and insightful for business organizations to confidently traverse in the right route if it is taken seriously. The continuous unearthing of integrated processes, platforms, patterns, practices and products are good indications for the bright days of big data phenomenon.

The implications of big data are vast and varied. The principal activity is to do a variety of tool-based and mathematically sound analyses on big data for instantaneously gaining big insights. It is a well-known fact that any organization having the innate ability to swiftly and succinctly leverage the accumulating data assets is bound to be successful in what they are operating, providing and aspiring. That is, besides instinctive decisions, informed decisions go a long way in shaping up and confidently steering organizations. Thus, just gathering data is no more useful but IT-enabled extraction of actionable insights in time out of those data assets serves well for the betterment of businesses. Analytics is

Related Content

A Deep Neural Network Model for Cross-Domain Sentiment Analysis

Suman Kumari, Basant Agarwaland Mamta Mittal (2022). *Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines (pp. 157-173).*

www.irma-international.org/chapter/a-deep-neural-network-model-for-cross-domain-sentiment-analysis/308485

Domain-Driven Data Mining: A Practical Methodology

Longbing Caoand Chengqi Zhang (2006). *International Journal of Data Warehousing and Mining (pp. 49-65).*

www.irma-international.org/article/domain-driven-data-mining/1774

A Methodology for Building XML Data Warehouses

Laura Irina Rusu, J. Wenny Rahayuand David Taniar (2005). *International Journal of Data Warehousing and Mining (pp. 23-48).*

www.irma-international.org/article/methodology-building-xml-data-warehouses/1750

Privacy Implications of Organizational Data Mining

Hamid R. Nemati, Charmion Brathwaiteand Kara Harrington (2004). *Organizational Data Mining: Leveraging Enterprise Data Resources for Optimal Performance (pp. 61-78).*

www.irma-international.org/chapter/privacy-implications-organizational-data-mining/27908

Multidimensional Business Benchmarking Analysis on Data Warehouses

Akiko Campbell, Xiangbo Mao, Jian Peiand Abdullah Al-Barakati (2017). *International Journal of Data Warehousing and Mining (pp. 51-75).*

www.irma-international.org/article/multidimensional-business-benchmarking-analysis-on-data-warehouses/173706