IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 56-58

The Recent Changes in Big Data Analysis

Prof. Prashant Dupate
M.Tech. (CSDP) (IIT KGP), MBA, M.Sc.(Stats), PGDCS&A

Abstract: *With the explosion of data sizes, the domain of big data is gaining enormous and prevalent popularity and research worldwide. The big data as well as big data repository possesses some peculiar attributes. Perhaps, analysis of big data is a common phenomenon in today's scenario and there are many approaches with positive aspects for this purpose. However, they lack the support to deal conceptual level. There are numerous challenges related to the performance of big data analysis. Precisely, these challenges are mainly related to enhance the effectiveness of big data analysis and optimum utilization of resources. Indeed, the lack of runtime level indicates the unawareness of various uncertainties pertaining to analysis. Furthermore, carefully analytic uncertainty is a challenging task and it can be created because of many reasons such as due to the fusion of huge amount of data from different sources. In addition, analytic uncertainty is also hard to predict the aspects for which the data is useful for the purpose of analysis. The main focus of this study is to illustrate different tools used for the analysis of big data in general.*

Keywords: *Hybrid Power Control System (HPCS), Automatic Transfer Switch (ATS), relays, contactors.*

I. Introduction

The amount of data generated today from all industry domains, also known as big data is huge, encompassing data gathering, data analysis, and data implementation process. Over the years, big data analytics trends are changing, from a departmental approach to business-driven data approach, embracing agile technologies and an increased focus on advanced analytics. Business enterprises need to implement the right data-driven big data analytics trends to stay ahead in the competition.

Previously, big data was primarily deployed by big businesses, who could afford the technology and channels used to collect and analyze the information. Today the scope of big data is changed leading to business enterprises large and small rely on big data for intelligent business insights. This has led to big data evolving at an unbelievably fast pace. The best example of the growth is big data in the cloud which has led to even small businesses taking advantage of the latest technology trends.

The never-ending stream of information is valuable to the business, but it can also be a challenge to draw actionable insights from a large data pool of data which may be unstructured. Even with these roadblocks, there's no denying the fact that big data offers business tremendous opportunities for growth. Here are the "Top 7 Big Data Analytics Trends" that will be the talk of the technology world in 2019 and beyond.

1. Fast Growing IoT Networks

Internet of Things (IoT) will be the trend, which will generate more than $300 billion annually by 2020. According to the latest industry trends and research reports, the global IoT market will grow at a CAGR of 28.5%. Business houses will rely on more data points to collect information for more detailed business insights.

IoT devices are a part of the larger concept of home automation, which can include lighting, heating and air conditioning, media and security systems.Long term benefits could include energy savings by automatically ensuring lights and electronics are turned off.

The IoT can assist in the integration of communications, control, and information processing across various transportation systems. Application of the IoT extends to all aspects of transportation systems (i.e. the vehicle, the infrastructure, and the driver or user). Dynamic interaction between these components of a transport system enables inter and intra vehicular communication, smart traffic control, smart parking, electronic toll collection systems, logistic and fleet management, vehicle control, safety and road assistance. In Logistics and Fleet Management for example, an IoT platform can continuously monitor the location and conditions of cargo and assets via wireless sensors and send specific alerts when management exceptions occur (delays, damages, thefts, etc.). This can only be possible with the IoT and its seamless connectivity among devices. Sensors such as GPS, Humidity, Temperature, send data to the IoT platform and then the data is analyzed and send further to the users. This way, users can track the real-time status of vehicles and can make appropriate decisions. If combined with Machine Learning then it also helps in reducing traffic accidents by introducing drowsiness alerts to drivers and providing self-driven cars too.

2. Predictive Analytics

Predictive Analytics offers customized insights that lead organizations to generate new customer responses or purchases and promote cross-sell opportunities. Predictive Analytics helps technology to integrate into diverse domains like finance, healthcare, automotive, aerospace, retailing, hospitality, pharmaceuticals, and manufacturing industries.

Predictive analytics is the practical result of Big Data and business intelligence (BI). What do you do when your business collects staggering volumes of new data? Today's business applications are raking in mountains of new customer, market, social listening, and real-time app, cloud, or product performance data. Predictive analytics is one way to leverage all of that information, gain tangible new insights, and stay ahead of the competition.

Organizations use predictive analytics in a variety of different ways, from predictive marketing and data mining to applying machine learning (ML) and artificial intelligence (AI) algorithms to optimize business processes and uncover new statistical patterns. It's basically computers learning from past behavior about how to do certain business processes better and deliver new insights into how your organization really functions. But before we get into all of the fascinating ways businesses and technology companies are employing predictive analytics to save time, save money, and gain an edge over the rest of the market, it's important to talk about exactly what predictive analytics is and what it's not.

3. Dark Data

Dark data in technology is the digital information that is currently not in use for business analysis. This data is acquired through various computer network operations which are not used in a manner to derive insights or for decision-making. As analytics and data become daily aspects of organizations, there is an increased need to understand that any data left unexplored is an opportunity lost and may lead to a potential security risk.

At the moment, we have solid evidence to suggest that as much as 90% of all data used in enterprises could be dark. Since industries are now storing large data volumes in the 'lake', it should be natural to tag the data appropriately as it gets stored. Perhaps the key is to extract the metadata out of this data and then storing it.

Profiling and exploring the data can be done using one or a combination of tools that are already available in the market. Cognitive computing and machine learning can further increase processing power and open up possibilities of making intelligent use of dark data.

Dark data may or may not have an identifiable structure. For example, most contacts and reports in organizations are structured. But over the course of time, they add up to the pile of dark data. Unstructured data can be small bits of personally identifiable info like birth dates and billing details. In the very recent past, this type of data would remain dark.

4. CDOs in Demand

The profile of the Chief Data Officer (CDO) has evolved and human resource personals are scouting for professionals who can fill this trendy job role. Though in demand, CDO is still a relatively new concept to many companies. Organizations have realized that they need to hire a CDO, so if you are a data leader managing enterprise-wide data cleaning, analysis, visualization and studying intelligent insights, CDO may be the work profile for you.

II. Methodology

Big Data fueled board and executive awareness in the potential power of data as a corporate asset, firms struggled with how best to organize around data -- as an activity, a business function, and a capability. Proponents argued the case for how critical data is as a tool for competitive advantage. In response, a set of industry leading firms coalesced around the realization that a new organizational role was needed, the equivalent of a 'data czar'. This new position has come to be known as the Chief Data Officer (CDO).

1**. Quantum Computing**

Tech giants like IBM, Microsoft, Google and Intel, race against each other to work rigorously in a bid to build the first quantum computer. Quantum Computing enables seamless data encryption, weather prediction, solving complex medical problems, real conversations and better financial modeling to make organizations develop quantum computing components, algorithms, applications and software tools on qubit cloud services.

We experience the benefits of classical computing every day. However, there are challenges that today's systems will never be able to solve. For problems above a certain size and complexity, we don't have enough computational power on Earth to tackle them.

To stand a chance at solving some of these problems, we need a new kind of computing. Universal quantum computers leverage the quantum mechanical phenomena of superposition and entanglement to create states that scale exponentially with number of qubits, or quantum bits.

Rather than store information using bits represented by 0s or 1s as conventional digital computers do, quantum computers use quantum bits, or qubits, to encode information as 0s, 1s, or both at the same time. This superposition of states—along with the other quantum mechanical phenomena of entanglement and tunneling—enables quantum computers to manipulate enormous combinations of states at once.

2. Open Source

2019 will witness more free data and software tools to become available on the cloud. Small organizations and start-ups alike will benefit the most of this data trend in 2019. Open source analytical languages like R, a GNU project associated with statistical computing and graphics has seen a huge adoption credit to the open source wave.

The term "open source" refers to something people can modify and share because its design is publicly accessible.

The term originated in the context of software development to designate a specific approach to creating computer programs. Today, however, "open source" designates a broader set of values—what we call "the open source way." Open source projects, products, or initiatives embrace and celebrate principles of open exchange, collaborative participation, rapid prototyping, transparency, meritocracy, and community-oriented development. Open source software is software with source code that anyone can inspect, modify, and enhance.

"Source code" is the part of software that most computer users don't ever see; it's the code computer programmers can manipulate to change how a piece of software—a "program" or "application"—works. Programmers who have access to a computer program's source code can improve that program by adding features to it or fixing parts that don't always work correctly.

3. Edge Computing

Edge Computing has been into the technological space streaming network performance for quite a while now. All credit to edge computing that data analytics is partly reliant on the network bandwidth to save data locally close to the data source. Edge Computing makes data to be handled and stored away from the silo setup closer to end users with processing taking place either in the device itself or in the fog layer or in the edge data center.

III. Performance Evaluation

No wonder businesses are increasingly turning to the edge to solve challenges in cloud infrastructure. Edge data centers work by bringing bandwidth-intensive content closer to the end user, and latency-sensitive applications closer to the data. Types of edge computing vary, including local devices, localized data centers, and regional data centers. But the objective is the same: to place computing power and storage capabilities directly on the edge of the network.

Edge computing also delivers business benefits to a wide variety of industries. Today's retailers rely on a 24/7 online presence to deliver superior customer experiences. Edge computing can prevent sites outages and increase availability for optimal site up-time. By ensuring factory floor operators stay connected to plant systems, edge computing can improve a manufacturer's operational efficiencies. And for healthcare practitioners, edge computing can place computing close to a device, such as a heart-rate monitor, ensuring reliable access to possibly life-saving, health-related data. The speed at which the world is generating data shows no signs of slowing down. As volumes mount, edge computing is becoming more than an IT necessity; it's a critical competitive advantage in the future.

IV. Conclusion

By ensuring factory floor operators stay connected to plant systems, edge computing can improve a manufacturer's operational efficiencies. And for healthcare practitioners, edge computing can place computing close to a device, such as a heart-rate monitor, ensuring reliable access to possibly life-saving, health-related data. The speed at which the world is generating data shows no signs of slowing down. As volumes mount, edge computing is becoming more than an IT necessity; it's a critical competitive advantage in the future.

References

[1]. Data Analytics Made Accessible, by A. Maheshwari
[2]. Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die by E. Siegel
[3]. Too Big to Ignore: The Business Case for Big Data, by award-winning author P. Simon
[4]. Analytics in a Big Data World: The Essential Guide to Data Science and its Applications, Bart Baesens
[5]. https://www.forbes.com/sites/bernardmarr/2017/06/06/the-9-best-free-online-big-data-and-data-science-courses/#3e1f5ed443cd
[6]. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0194-3