

Big Data Applications

Gurwinder Singh

Assistant Professor, Department of Computer Science, , Sikh National College Banga SBS

Nagar, Punjab, (India)

ABSTRACT

Big Data plays a vital role in various fields. Big Data nowadays being a challenging research area has attracted a lot of attention from academia, industry as well as government. Big Data is collection of large and complex data sets that are difficult to process using traditional data processing tools. We are creating trillions of data all over the world. This data is coming from social networking sites, scientific experiments, mobile conversations, sensor networks and various other sources. These day's Big data is most preferably used in enterprises, organizations, companies and business etc...So Big data has number of applications in various fields such as banking, agriculture, chemistry, data mining, cloudcomputing, finance, marketing, stocks, BDA, health care etc...This paper explanation some of applications of Big Data. Various papers based on big data have been studied so far and I explore the usage of Big Data in several areas like data mining, cloud computing, Banking, Marketing, Health care, Finance, Enterprise applications.

Keywords— Big Data, Big Data Analytics, Big data Applications.

I. INTRODUCTION

1.1. Big Data

The term Big Data describes large amount of data that is structured, semi-structured and unstructured which has potential to be mined for information. There is no defined boundaries of big data in terms of quantity, but this term is often used while talking about the petabytes and Exabyte of data.

1.2. Big Data Analytics

The process of examining large data sets that containing a variety of data types can be termed as Big Data Analytics. These analytical findings can lead to more effective marketing, new revenue opportunities, better customer service, improved operational efficiency, competitive advantages over rival organizations and other business benefits. Basic goal of big data analytics is to help companies make more informative business decisions by enabling data scientists, predictive modellers and other analytics professionals to analyse large volumes of transactional data, as well as other forms of data that may be untapped by more conventional Business Intelligence(BI) programs. This may include web server logs and Internet click stream data, social media content and social network activity reports, text from customer emails and survey responses, mobile phone call detail records and machine data captured by sensors and connected to the Internet of Things.

The term Big data coined in the first decade of the 21st century, and the first organization to clinch it were online and start-up firms. Google, LinkedIn, and eBay and Face book were built around big data from the

beginning. They did not have to join or integrate big data with more traditional sources of data and the analytics performed upon them, because they didn't have that much of traditional forms. They didn't have to merge big data technologies with their traditional IT infrastructures because these infrastructures didn't exist.

1.3.Hadoop and No SQL Free Software:

Hadoop is an open-source distributed file system which is capable of storing and processing large volumes of data in parallel using commodity hardware. Hadoop came from companies like Google and Yahoo, who needed a cost effective way to build search indexes. Engineers of large companies knew that traditional databases would be expensive and technically unwieldy, so they came up with an alternative that they built themselves. They gave it to the Apache Software Foundation so others could benefit from their innovations.

At present, many companies are implementing Hadoop software from Apache as well as third-party providers such as Cloud era, Horton works, EMC, and IBM. Developers see Hadoop as a cost effective way to get their arms around large volumes of data. Companies are using Hadoop to process, store and analyse large volumes of Web log data so they can get a better feel for the browsing and shopping behaviour of their customers.

II.ROLE OF BIG DATA

1) *In Enterprise:* For enterprises around the world, in many industries and in database, analytics are providing a competitive advantage. When data doesn't have to commute to work and back, it can deliver faster insights that help business people make informed decisions in realtime for less expense than traditional data analysis tools.

2) *In BDA:* Big Data Analytics Applications (BDA Apps) are a new type of software applications, which analyse big data using massive parallel processing frameworks (e.g., Hadoop). Developers of such applications typically develop them using a small sample of data in a pseudo-cloud environment. Afterwards, they deploy the applications in a large-scale cloud environment with considerably more processing power and larger input data.

Big Data Analytics Applications (BDA Apps) are a new category of software applications that leverage largescale data, which is typically too large to fit in memory or even on one hard drive, to uncover actionable knowledge

using large scale parallel-processing infrastructures [1]. The big data can come from sources such as runtime information about traffic, tweets during the Olympic Games, stock market updates, usage information of an online game [2], or the data from any other rapidly growing data-intensive software system.

3) *In Stock:* A private stock exchange in Asia uses in database analytics to establish a comprehensive system to detect abusive trading patterns to detect fraud.

4) *In Clustering:* Using clustering (K-means algorithm) through a simple point and click dialog, users can automatically find groups within data based on specific data dimensions. With clustering, it is then simple to identify and address groups by customer type, text documents, products, patient records, click path, behaviour, purchasing patterns, etc.

5) *In Data Mining:* Decision Tree--Datameer's decision trees automatically help users understand what combination of data attributes result in a desired outcome. Decision trees illustrate the strengths of relationships

and dependencies within data and are often used to determine what common attributes influence outcomes such as disease risk, fraud risk, purchases and online signups. The structure of the decision tree reflects the structure that is possibly hidden in your data.

6) *In Banking:* The use of customer data invariably raises privacy issues. By uncovering hidden connections between seemingly unrelated pieces of data, big data analytics could potentially reveal sensitive personal information. Research indicates that 62% of bankers are cautious in their use of big data due to privacy issues. Further, outsourcing of data analysis activities or distribution of customer data across departments for the generation of richer insights also amplifies security risks.

7) *In Sap:* Sybase (now SAP) laid the groundwork for the analytical platform market when it launched the first columnar database in 1995. Tera data was also an early forerunner, shipping the first analytical appliance in the early 1980s. Netezza kicked the current market into high gear in 2003 when it unveiled a popular analytical appliance and was soon followed by dozens of start-ups. Recognizing the opportunity, all the big names in software and hardware. They are Oracle, IBM, HP, and SAP subsequently jumped into the market, either by building or buying technology, to provide purpose-built analytical systems to new and existing customers.

8) *In Credit Cards:* Credit card companies rely on the speed and accuracy of in-database analytics to identify possible fraudulent transactions. By storing years' worth of usage data, they can flag atypical amounts, locations, and retailers, and follow up with cardholders before authorizing suspicious activity.

9) *In Telecom:* In telecom big data plays a very good role. Service providers are trying to compete in the cutthroat world of telecom services. Where more and more subscribers rely on over-the-top (OTT) players as providers of value-added services are focused on increasing revenue, reducing opex, chum and enhancing the customer experience as key business objectives. Operators believe that big data and advanced analytics will play a critical role in helping them meet their business objectives. In the same survey, respondents indicate critical use case scenarios in the context of big data and advanced

analytics where they are investing now and where they plan to invest in the next three years. Operators face an uphill challenge when they need to deliver new, compelling, revenue generating services without overloading their networks and keeping their running costs under control. The market demands new set of data management and analysis capabilities that can help service providers make accurate decisions by taking into account customer, network context and other critical aspects of their businesses. Most of these decisions must be made in real time, placing additional pressure on the operators. Real-time predictive analytics can help leverage the data that resides in their multitude systems, make it immediately accessible and help correlate that data to generate insight that can help them drive their business forward.

10) *In Consumer Goods:* A maker of consumer products collects consumer preference and purchasing data extracted from surveys, purchases, web logs, product reviews from online retailers, phone conversations with customer call centres, even raw text picked up from around the Web. Their ambitious goal: to collect everything being said and communicated publicly about their products and extract meaning from it. By doing this, the company develops a nuanced understanding of why certain products succeed and why others fail. They can spot trends that can help them feature the right products in the right marketing media.

11) Where to Use Hadoop: In every vertical there are data tasks with which Hadoop can assist. These tasks have different terms depending on the industry but they all come down to either advanced analytics or data processing.

12) In Health care: Traditionally, the health care

industry has lagged behind other industries in the use of big data, part of the problem stems from resistance to change providers are accustomed to making treatment decisions independently, using their own clinical judgment, rather than relying on protocols based on big data.

Many health care stakeholders have under invested in information technology because of uncertain returns. Although their older systems are functional, they have limited ability to standardize and consolidate data. The nature of health care industry itself also creates challenges: while there are many players, there is no way to easily share data among different providers or facilities, partly because of privacy concerns. Even within a single hospital, payor, or pharmaceutical company, important information often remains siloed within one group or department because organizations lack procedures for integrating data and communicating findings.

Health care stakeholders now have access to promising new threads of knowledge. This information is a form of "big data," so called not only for its sheer volume but for its complexity, diversity, and timelines.

Pharmaceutical industry

exports, payers, and providers are now beginning

to analyse big data to obtain insights. Although these efforts are still in their early stages, they could collectively help the industry address problems related to variability in health care quality and escalating health care spend. Researchers can mine the data to see what treatment are more effective for particular conditions, identify patterns related to drug side effects or hospital read missions, and gains other important information that can help patients and

reduce costs. Recent technologic advances in the industry have improved their ability to work with such data, even though the files are enormous and often have different database structures and technical characteristics.

13) In Agriculture: A biotechnology firm uses sensor data to optimize crop efficiency. It plants test crops and runs simulations to measure how plants react to various changes in condition. Its data environment constantly adjusts to changes in the attributes of various data it collects, including temperature, water levels, soil composition, growth, output, and gene sequencing of each plant in the test bed. These simulations allow it to discover the optimal environmental conditions for specific gene types.

14) In Finance: A major financial institution grew wary of using third-party credit scoring when evaluating new credit applications. Today the bank performs its own credit score analysis for existing customers using a wide range of data, including checking, savings, credit cards, mortgages, and investment data.

15) In Economy: Designed from the ground up to deal intelligently with commodity hardware, Hadoop can help organizations transition to low-cost servers.

16) In Conservation: Keeping data in a merged, isolated system provides business intelligence benefits and is both financially and ecologically sound.

17) In Marketing: Marketers have begun to use facial recognition software to learn how well their advertising succeeds or fails at stimulating interest in their products. A recent study published in the Harvard Business

Review looked at what kinds of advertisements compelled viewers to continue watching and what turned viewers off. Among their tools was “a system that analyses facial expressions to reveal what viewers are feeling.” The research was designed to discover what kinds of promotions induced watchers to share the ads with their social network, helping marketers create ads most likely to “go viral” and improve sales.

III. CONCLUSION

Big Data is a powerful tool which makes things ease in various fields as said above. Big data used in number of applications like banking, agriculture, chemistry, data mining, cloud computing, finance, marketing, stocks, BDA, health care etc...

REFERENCES

- [1] D. Fisher, R. DeLine, M. Czerwinski, and S. Drucker, “Interactions with big data analytics,” *interactions*, vol. 19, no. 3, pp. 50–59, May 2012
- [2] N. Wingfield, “Virtual product, real profits: Players spend on zynga’s games, but quality turns some off,” *Wall Street Journal*.
- [3] Brandon Bunker, Senior Director of Customer Analytics and Intelligence, Vivint, “Big Data Insights Platform for Rapid Data Discovery,” 2013.
- [4] Weiyi Shangy, Zhen Ming Jiangy, Hadi Hemmatiy, Bram Adamsz, Ahmed E. Hassany, Patrick Martinx, “Assisting Developers of BigData Analytics Applications When Deploying on Hadoop Clouds” *Database Systems Laboratory, School of Computing, Queen’s University, Kingston, Canada*.
- [5] Ari Banerjee senior analyst, heavy reading, “Big data and advanced analytics in Telecom: A Multi-Billion-Dollar Revenue Opportunity,” December 2013.
- [6] Thomas H. Davenport, Jill Dyche, “Big Data in Big Companies,” in *International Institute for Analytics* May 2013.
- [7] By Steve Lucas, Executive Vice President and General Manager, Database and Technology, SAP, “Big Data Analytics Guide 2012.
- [8] Formerly Booz & company, “Benefiting from big data: A new approach for the telecom industry,” in 2013.
- [9] Peter Groves, Basel Kayyali, David Knott, Steve Van Kuiken, “The big data revolution in health care,” enter for US Health System Reform Business Technology Office, published in January 2013.
- [10] “Data-driven healthcare organizations use big data analytics for big gains” by IBM software.
- [11] “Deep learning applications and challenges in big data analytics” by Maryam M Najafabadi, Flavio Villanustre, Taghi M Khoshgoftaar, Naeem Seliya, Randall Wald and Edin Muharemagic.
- [12] “Big data is the future of Healthcare” by Cognizant 20-20 insights September 2012.
- [13] “Big Data: Trends, Strategies, and SAP Technology” by Carl W.Olofson, Dan Vesset August 2012.