

# Big Data Analytics using Digital 2.0 Transformation

Abha Chaudhary  
DYPSOET

SavitribaiPhule Pune University  
Maharashtra, India

PaulomiNath  
DYPSOET

SavitribaiPhule Pune  
University  
Maharashtra, India

Nidhi Wakchoure  
DYPSOET

SavitribaiPhule Pune University  
Maharashtra, India

NilkamalGotarne  
DYPSOET

SavitribaiPhule Pune University  
Maharashtra, India

BhagyashreeDhakulkar  
DYPSOET

SavitribaiPhule Pune University  
Maharashtra, India

## ABSTRACT

The tools and technology used in organizations and enterprises to improve analytics method has been described. Various developing organizations, enterprises and companies yield tremendous amount data everyday. Companies require to understand customer's necessities from the information available to it. This paper explains the tools and technologies which can help in analyzing the data by finding hidden patterns behind it.

## General Terms

Big Data Analytics, Digital 2.0

## Keywords

Big data, analytics, Microsoft power BI, machine learning, dashboard, predictive analysis

## 1. INTRODUCTION

Previously, the biggest drawback of an organization came from the inability in the prediction of customer's requirement [2]. The reason behind this inability was due to lack of analysis process as there was no technology present for linking the datasets from different data centers. As a result of these circumstances organizations were unable, to depict the fault and errors caused by previously undertaken step within a process, to improve company's performance by taking strategic decision [1]. Later on, after evolution of analytics, the organization started better understanding their clients need and goal.

To encourage organization and its performance in an appropriate way utilization of big data analytics [3], became as important as analytics. The availability of massive amount of data[4], which becomes difficult to store, process and mine using conventional database, primarily because of huge, complex, unstructured data becomes big data [9]. Big data analytics is important in order to obtain effective data context by focusing on accurate information. And also tells what should be the primary aspect of the business. Online firms like Google, eBay, facebook, LinkedIn, were the first to accept the concept of big data. Business oriented specialized tools and technologies have been introduced to intensify big data analytics.

## 2. PRESENTLY USED TECHNOLOGIES

### 2.1 Analytics

It is the process of finding different patterns hidden inside the data to form statistical view. It drills inside the data and finds

the information that can be useful in different aspects of organization.

### 2.2 Big data

Different organizations, Firms, enterprises, social networking sites, other websites produce data in large amount everyday, which contains metadata (i.e. data about data) this forms big data. For example, facebook approximately produces 500+ terabytes of data everyday. Since it is in tremendous amount, it may contain different patterns which help in analytics.

### 2.3 Big Data Analytics

Nowadays deluge data is generated by different firms (i.e., healthcare, companies, internet, social sites etc.). Data from all this domains forms a big data it is not efficient to manage such a large amount of data. There are many tools to manage this data but sometimes it is unmanageable [10], so big data analytics is mostly used for this purpose. By using big data analytics different firms can use the previous and present data and forecast the future. Mostly in organizations the company can use the past and present response of the customer and depending upon these values future plans can be made and these can lead to the profit of organization and also minimization of unwanted data.

#### 2.3.1 CRISP-DM

It is one of the way for performing data analytics [3], it consist of six phases that are:

1. Problem/business understanding determines the goal of the organization and the plan by which the organization can gain lots of profit.
2. Data understanding leads to study of the previous and present data.
3. Data preparation includes the collection of data which is important for the organization based on customer requirements and reviews.
4. Modelling leads to selection of suitable model and test designs.
5. Evaluation includes comparing the current model with the initial success plans.
6. Deployment performs the analytics on regular basis and monitors the data.

Big data analytics makes data more digitized and also helps number of organizations to predict the future plans based on their past and present data related to customer actions and

reviews. Mostly organizations aims that their stakeholders should be increased and also the value of organization should be increased for this purpose data analytics is used by numerous organizations.

## 2.4 Machine Learning

Machine learning is closely related to 'Artificial Intelligence and Big data (possibility of availability at different sources)'. Informal definition of machine learning [5] can be machines learning from existing data and making prediction on the data, predictions can be suggestions as well as decisions. Machine learning Can be defined as the ability of computers to understand and learn things without being externally programmed. "Can machines think? Can machine perform the tasks of humans?" questions like this gave rise to machine learning.

For a computer to perform a certain specific task requires a programmer to create a program to complete the specified task, this activity requires time and efforts. Making computer to perform such tasks of human and interaction with the human at the same time can me made possible using machine learning. Real time examples of machine learning are spam recognition, content filtering, optical character recognition, search engine, advertisement blocker for browser, credit scoring, fraud detection, stock trading and many more.

### 2.4.1 Classification of Machine Learning:

1. Supervised learning: Problem statement is given to the computer by a supervisor [6], [7] with their solutions the goal is to make the computer learn the general mapping of inputs to outputs.
2. Unsupervised learning: No labels and supervisor is given to the computer, it has to be on its own to map structure in its input, goal is to find the hidden pattern in data.
3. Reinforcement learning: Interaction of a computer program with live environment and explicitly informing the program about the goals. A good example is robotics.

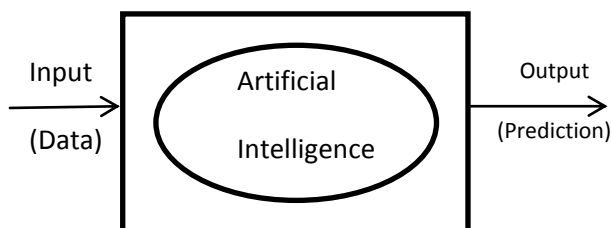


Fig 1: Machine Learning

A good live example of machine learning is, in 2006 the online movie company Netflix organized a competition to find a program that will help them to predict user preferences and improve accuracy in recommendations. A team of computer programmers won 1 million dollars in 2009 for the same. The main approach of machine learning is to make computers more automated, make them do what humans are capable to do i.e. think and act like humans resulting reduced human interaction which will simplify tasks and save time and efforts with the help of artificial intelligence.

## 3. TOOLS

### 3.1 Dashboard

Dashboards are also called as Management information systems. These are linked with the databases to create reports. It gives an overview about the reports to focus on the metrics, an enterprise care about. It is a user interface that represents the information of particular data in such a manner which is easy to read. In an organization and enterprise, dashboard provides view of key performance indicators relevant to any business process. Dashboards tell the user where the maintenance is required and shows the summaries, comparisons in graphical form. This helps the organization in making strategic decision to improve company's performance and to fulfill customer's need.

#### 3.1.1 Key properties of Dashboard:

1. Transparency
2. User defined
3. Control over data
4. Provides meaningful report

Every dashboard serves at distinct purpose, it depends on user's choice which dashboard solves the problem effectively. Dashboards can be divided into following categories:

#### 3.1.2 Strategic

These dashboards provide quick overview of the data to the user at all level, to make strategic decision which can check the ongoing performance of a business and health of an organization. This further helps in finding fault and to increase the performance.

#### 3.1.3 Analytical

These dashboards provide compared information of data across time and multiple variables. Also provides view of data in increasing detail, compares data at different instants of time and gives the analyzed information after drilling the data. Since, data stored in each second is analyzed, it is important to maintain the ability of comparison. If the ability is lost analytical dashboard will no longer be helpful to accomplish goal of user.

#### 3.1.4 Operational

Data analysis required monitoring of activities and events that constantly change; sometimes it may require attention at that moment, in that case operational dashboard works. It monitors the real-time operations on the data and alerts users at any deviation from the path or if any fault occurs. They provide exact information to the user to recover from fault or deviation.

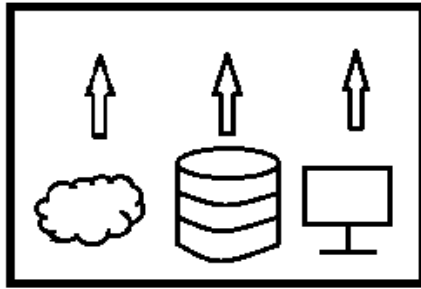
### 3.2 Microsoft Power BI

Data centers store data in large number; prior to technology linking of this data was not possible. Microsoft Power BI is a tool which solves this problem. It collects organizational data, organizes and transforms into rich perceptible information, to focus on what matters to the enterprise. It creates dashboards which help in detecting faults for further improvement.

#### 3.2.1 Phases involved in the working of Power BI:

##### Step 1: Collect the data

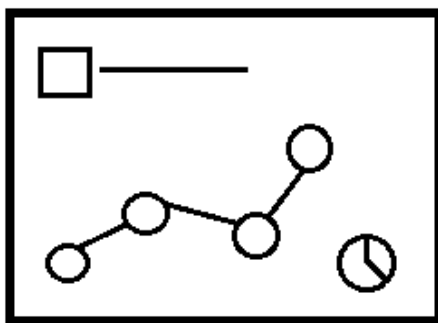
It collects the interested data from different data centers, stores it into excel spreadsheets. This saves user's time instead of searching in different directions and direct starts exploring.



**Fig 2: Collect the data**

*Step 2: Create data visuals*

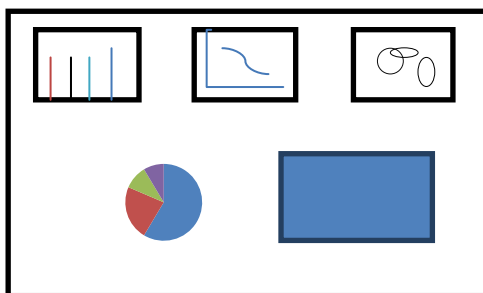
It creates the data visuals by using different gestures like line, circles etc. This data visuals can be easy to understand and percept the underlying data behind it.



**Fig 3: Create data visuals**

*Step 3: Collect and Share*

It collects all the visuals created and generate dashboards. It creates an integrated view of information that can help the team to take immediate decisions and to improve the performance of ongoing process.



**Fig 4: Collect and share**

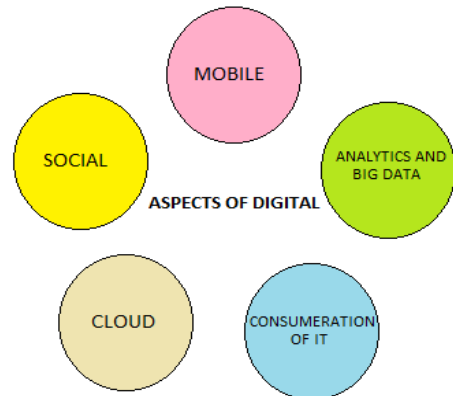
### 3.3 Digital 2.0 Transformation

#### 3.3.1 Concept of Digital and Transformation:-

The Electronic Technology which generates the data, Stores it and processes it among two states i.e. Positive and Negative is termed as digital. The number 1 represents the Positive Expression while the number 0 represents the Negative expression. The digits are described as a byte. The Strings of the expression 0 and 1 transmits and store the particular data using digital technology.

Formerly the data transmission was restricted to analogue technology. In this technology the data being conveyed as electronic signals of varying amplitude and frequency which is being added to carrier waves. This technology has been used in phone transmission and broadcasting. Afterward improvised technology came and the digital technology evolved. Primarily it is used with new physical media like the fiber optic and satellite transmission. Some interesting technologies evolved from digital technologies are:-Digital photos, Digital video, Digital fingerprints, Automated telephone operators. Digital technology became one of the most influential and major technology now.

#### 3.3.2 Five Major Aspects of Digital 2.0:



**Fig 5: Aspects of Digital 2.0**

One of the foundational components of business is Transformation. The transformation consider every function of business-What is the operations and technology a company is using, sales, marketing, what it is selling to the customer and how it is going to the market.

The business goes through transformation when it has been fail to evolve. It is not being transformed by its own choice without the necessity.

#### 3.3.3 Digital transformation

The modification which is associated with the applications of digital technology in aspects of human society is termed as digital transformation. Shifting of organization to the new ways of working and thinking using digital, social, mobile and other promising technologies from a heritage approach is the process of digital transformation Expectations of how lives can be made easier and more efficient are accelerated by digitization. The devices that is being used is expected to perform related engaging content automatically.

#### 3.3.4 Role played by different Organization in digital transformation

Many Global Consultancies like Accenture, Pwc manages to transform program on behalf of large businesses. It takes a huge amount of time, huge amount of resources and business skills which are found in such global consultancies for transforming a large organization.

On the other hand creative agencies target in helping the business to evade the necessity of digital transformation. These agencies are expertise in understanding the need of consumers, technologies and markets and what further going to happen. Working with their client let them to understanding what is coming and helping the client to evolve to meet it.

### 3.3.5 Digital 2.0

As there is enhancing competition across all sectors, the companies need to steadily adopt their services to the customer's requirements. The working methods and processes in business need to take on concept of incessant improvement, transversal and openness.

Enterprise 2.0 helps the employees to team up, allocate and classify information via web 2.0 making it the ultimate tool to enhance ingenuity and responsiveness. Enterprise 2.0 aims to promote communication throughout organizations by the use of shared tools and social platforms. These tools conduct to intense changes in working methods:-

- The contributors are users and their information contributes as a whole to the organization's development..
- Profile of the employees and searches against those profiles publishing can encourage unexpected and unplanned associations for solving business problems and point expert on different topics

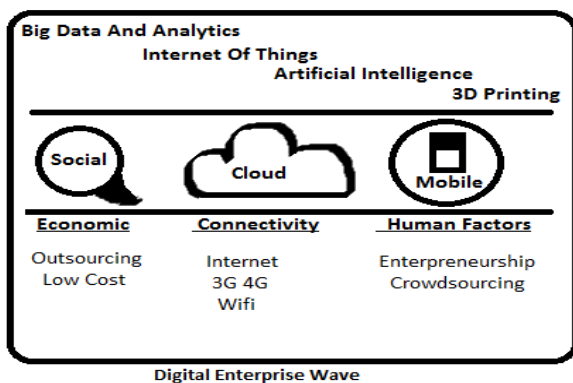


Fig 6: Digital Enterprise Wave

## 4. PRESENT SCENARIO

Data analytics provides acceleration and improvement in the process of genuine selection of investments. As an example, Google+ is one of the integral parts of Google. Any user who performs activity on Google+ gets stored in the Google database. Google data analytics examines this data and finds the hidden information behind it. As a result of this Google updates its information every minute.

## 5. FUTURE ASPECTS

Survey of large data sets which is big data will become a fundamental basis of computation, supporting new waves of productivity growth and modernization of individual firms. The escalating volume and thorough information captured by enterprises, rise of multimedia services, social media like facebook, whatsapp, LinkedIn etc., and the IoT will ignite exponential growth in data for the predictable future.

## 6. CONCLUSION

In this paper, a brief description of tools and technologies is outlined for data analytics. Utilization of these mentioned tools has been initiated by many firms and still going on. These will be beneficial in future generations as it gives the output in favor of both the customer and enterprise. Surveying large data sets will become easier for enterprises. Digitized technologies will accelerate productivity growth of firms and organization.

## 7. REFERENCES

- [1] Qiufen Xia, Weifa Liang and Zichuan Xu, "Data Locality-Aware Query Evaluation for Big Data Analytics in Distributed Clouds", 2014 Second International Conference on Advanced Cloud and Big Data.
- [2] Ankur Balar, Nikita Malviya, Swadesh Prasad, Ajinkya Gangurde, Forecasting consumer behavior with innovative value proposition for organizations using Big Data Analytics, 2013 IEEE International Conference on Computational Intelligence and Computing Research
- [3] Parth Chandarana M. Vijayalakshmi, "Big Data Analytics Frameworks", 2014 International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA)
- [4] M. Riedel, A.S. Memon and M.S. Memon, "High Productivity Data Processing Analytics Methods with Applications", MIPRO 2014, 26-30 May 2014, Opatija, Croatia
- [5] Nicolas Bredeche, Zhongzhi Shi, Senior Member, IEEE, and Jean-Daniel Zucker, "Perceptual Learning and Abstraction in Machine Learning: An Application to Autonomous Robotics", IEEE March 2006
- [6] Li Cunhe, Wu Chenggang, "A New Semi-Supervised Support Vector Machine Learning Algorithm Based on Active Learning", 2010 2nd International Conference on Future Computer and Communication
- [7] Xiaofeng Liao, Liping Ding\*, Yongji Wang\*†, "Secure Machine Learning, A Brief Overview", IEEE Computer Society 2011.
- [8] Andréa Matsunaga and José Fortes, "On the use of machine learning to predict the time and resources consumed by applications", 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing
- [9] Peter Augustine, Leveraging Big Data Analytics and Hadoop in Developing India's Healthcare Services, International Journal of Computer Applications (0975 – 8887) Volume 89 – No 16, March 2014
- [10] Rajeshwari.D, "State of the Art of Big Data Analytics: A Survey", International Journal of Computer Applications (0975 – 8887) Volume 120 – No.22, June 2015.