https://doi.org/10.46335/IJIES.2024.9.4.7 e-ISSN: 2456-3463
Vol. 9, No. 4, 2024, PP. 35-42

International Journal of Innovations in Engineering and Science, www.ijies.net

Role of Big Data Analytics In Industry 4.0

Bhagyashri S. Patil

GF's GCOE, Jalgaon

*E-Mail:**patilbhagyashri1693@gmail.com***

***Received on**: 19 April, 2024* ***Revised on**: 21 May, 2024* ***Published on**: 22 May, 2024*

***Abstract –** Industry 4.0 is a term used for the fourth industrial revolution: the digitization and automation of manufacturing. Advances in networking, machine learning, data analytics, robotics, 3D printing, and other technologies are making vast improvements on industrial processes and reducing our dependence on human labor and decision-making. By leaning into digital solutions, manufacturing can reduce human error, shorten time to market, and increase the speed at which industrial processes can adapt to new information. Now a day, big data in industry 4.0 is a major concern for current research for the organizations that are motivated to invest in these kinds of projects. The big data are known as the large quantity of data collected from various resources that potentially could be analyzed and provide valuable insights and patterns. In industry 4.0 the production of data is massive, and thus, provides the basis for analysis and important information extraction.*

The paper first introduces two major areas Industry 4.0 and Big Data Analytics (BDA) with its concepts, characteristics and processes. The main aim of the paper is to provide the impact of big data analytics in industry 4.0 environments by providing both a positive and a negative perspective of the subject.

Keywords- Industry 4.0, Big Data Analytics, 5 V's of big data.

I -BACKGROUND OF INDUSTRY 4.0

The term "Industry 4.0" is used to signify the beginning of the fourth industrial revolution which is based on the digitization and automation of manufacturing. Today, it is in the midst of a fundamental shift in the way products are produced, and it's deeply tied to the future of the Internet of Things (IOT).

Now a days various technological advances are knocking the door of industrial processes and reduces our dependencies on human labor and decision making. It includes, advances in networking, machine learning, data analytics, robotics, 3D printing and many more. By leaning into digital solutions, manufacturing can reduce human error, shorten time to market, and increase the speed at which industrial processes can adapt to new information. The following figure 1shows the industrial revolution which includes four phases of revolution [1][2].

Figure 1: Industry revolution

The first industrial revolution is actually just of four revolutions so far.

- The first revolution moved from manual to mechanical by harnessing steam and water power, replacing raw muscle. It really was the revolution, and, one must thanks to invention of steam machines, the usage of water and steam

https://doi.org/10.46335/IJIES.2024.9.4.7 **e-ISSN: 2456-3463**
Vol. 9, No. 4, 2024, PP. 35-42

power and all sorts of other machines. It would lead to the industrial transformation of society with trains, mechanization of manufacturing and loads of smog.

- The second revolution took advantage of electricity and the assembly line to generate mass production. It is typically seen as the period where electricity and new manufacturing 'inventions' which it enabled, such as the assembly line, led to the area of mass production and to some extent to automation.
- The third revolution took off using electronics and software, the most iconic example of these two coming together being the computer—and the Programmable Logic Controller (PLC) in the industrial setting. This gave rise to an era of high-level automation, it had everything to do with the rise of computers, computer networks (WAN, LAN, MAN,…), the rise of robotics in manufacturing, connectivity and obviously the birth of the Internet, that big game changer in the ways information is handled and shared, and the evolutions to e-anything versions of previously brick and mortar environments only, with far more automation.
- The fourth revolution is the revolution happening right now, according to this study. The physical world, the digital world, and the virtual world are colliding together, creating smart products, procedures and processes, and smart factories. The study goes on to describe this revolution being brought on primarily by **cyber-physical systems and the Internet of Things** to create a highly intelligent, integrated, and automated manufacturing ecosystem that spans far beyond the traditional "factory floor."

Cyber-physical systems, refers to the integration of physical processes, computation, and networking. It is the integration of embedded computers and networks monitor and control the physical processes with feedback loops, where physical processes affect computations and vice versa. Whereas **Internet of Things (IOT)** refers to an interconnected network of device-to-device communication that does not need a human or computer interface, usually taking advantage of the Internet [1][2].

- **Industry 4.0 (1.1)**

Industry 4.0 is not merely a matter of connecting machines and products thought the Internet. But it is a **paradigm shift** in way to organize, manage, and approach business. As a society, people are starting to feel the impacts of Industry 4.0 already. Not only are companies investing, but also governments around the world are pouring a lot of money into this idea as the way of the future. The working group characterized Industry 4.0 as a concept that is focused on creating smart products, procedures and processes, and smart factories. But the statement is so grand and unclear.

At its most basic level, Industry 4.0 is a broad term that encompasses different perspectives, industries, corporate functions, technologies, and fields. The focus and understanding of Industry 4.0 are constantly evolving due to the high level of activity and continual development of new approaches, concepts, and solutions on the part of solution providers, hardware manufacturers, businesses, research institutions, and even governmental agencies. Industry 4.0 is characterized by, among others [3],

- even more automation than in the third industrial revolution,
- the bridging of the physical and digital world through cyber-physical systems, enabled by Industrial IOT,
- a shift from a central industrial control system to one where smart products define the production steps,
- closed-loop data models and control systems and
- Personalization /customization of products.

The goal is to enable autonomous decision-making processes, monitor assets and processes in real-time, and enables equally real-time connected value creation networks through early involvement of stakeholders, and vertical and horizontal integration.

2.1 Industry 4.0 Outcomes

Industry 4.0 is having an international impact, in terms of connectivity across continents and the way it's transforming our global economies. Following are five of the biggest and most noteworthy benefits manufacturers can expect from Industry 4.0 [4].

- **Optimized Processes**

All of the Industry 4.0 connectivity – sensors, IOT, AI, etc. – services one primary purpose: optimizing manufacturing processes. Automation allows manufacturers to work faster, data analytics empowers leadership to make data-driven decisions to increase efficiency, predictive maintenance means less downtime for machines, and monitoring systems

https://doi.org/10.46335/IJIES.2024.9.4.7 e-ISSN: 2456-3463
Vol. 9, No. 4, 2024, PP. 35-42

provide real-time yield optimization across the operation.

This optimized processes and maximized efficiency means a lot to the manufacturer- it translates to revenue increases and improved customer service. When manufacturers are able to get the most out of their production with sensor-monitored machines, all while giving personalized attention and fast service to customers via AI and field service, and can see true benefits of the connected factory.

- **Greater Asset Utilization**

Industry 4.0 allows for greater flexibility across the manufacturing operation, which translates to better asset utilization and therefore a potential for revenue increases. With the automation – autonomous mobile robots (AMR) can handle menial tasks such as product transportation, leaving skilled human workers to do more higher-value tasks.

- **Higher Labor Productivity**

When employees feel more secure on the job, they can focus better and accomplish more tasks during the day. Worker safety is one of the biggest benefits of IOT solutions on the manufacturing floor – sensors on site and worn by workers are monitored constantly to ensure a safe and healthy work environment.

Industry 4.0 is also expanding the skills range of many manufacturing workers. As new technologies come into the operation, workers are learning new skills that improve operational efficiency and their skillset. Think of cobots (collaborative robots) – people and robots working alongside each other in manufacturing workflows, maximizing efficiencies and revenues.

- **Supply Chain and Inventory**

IOT-enables sensors and data analytics give manufacturers insight into the entire supply chain and production process. This level of visibility combined with AI and machine learning capabilities means that supply chain optimization can be achieved in real-time.

It can be called as Supply Chain 4.0, defined as "the application of the Internet of Things, the use of advanced robotics, and the application of advanced analytics of big data in supply chain management: place sensors in everything, create networks everywhere, automate anything, and analyze everything to significantly improve performance and customer satisfaction."

- **After-Sales Services**

The predictive analytics, virtual reality, and remote monitoring that are pillars of Industry 4.0 also translate to the consumer space after manufacturing. While this doesn't directly impact a manufacturer, if they create goods that are capable of IOT connectivity, they can drastically improve customer and field service offerings. With connected devices, manufacturers can monitor product performance, scheduling maintenance before an issue arises and therefore preventing any sort of customer dissatisfaction.

2.2. Key concepts of Industry 4.0

The challenge with defining Industry 4.0 is that it encompasses such a broad array of technologies. To understand what makes the fourth industrial revolution and to recognize the distinct qualities that these various concepts bring into the manufacturing and how it completely change the way we produce goods. Here are some of the core concepts and technologies that fulfill the design principles of Industry 4.0 [2][4].

- **Internet of Things (IOT)**
  Networking devices and retrieving data from equipment is a foundational part of Industry 4.0. IOT isn't new, but it *is* rapidly expanding. Advances in cellular technology like 5G and specialized IOT networks like LTE-M and NB-IOT are opening the door to new cellular capabilities so one can concentrate more devices in the same space, and build applications that solve more complex problems.
- **Cyber Physical Systems (CPS)**
  Cyber Physical Systems tie together robotics, IOT, and machine learning. It is essentially any mechanical process that is automatically controlled by software. Using sensors and other inputs from the mechanical components, the software runs algorithms that determine how it should control the machinery, equipment, or infrastructure. A CPS can build to be working in heterogeneous environment, which makes it easy to fulfill changing needs of Manufactures.
- **Big Data analytics**
  "Big Data" refers to data sets that are exceptionally large or complex, making them difficult to process in a timely manner—even with software. A single IOT device can produce enormous quantities of data to analyze. When you have tens of thousands of interconnected devices, the challenge is to find the most valuable ways to collect, store, analyze, and use that data. In Industry 4.0, wrestling with Big Data is the key to implementing predictive maintenance and understanding user behavior.

https://doi.org/10.46335/IJIES.2024.9.4.7 e-ISSN: 2456-3463
Vol. 9, No. 4, 2024, PP. 35-42

- **Horizontal and vertical systems integration**

  In Industry 4.0,every company, process, employee, department, and piece of equipment involved in manufacturing needs to communicate. From research and development, to the supply chain, to the production floor, to customer service, to marketing, and to sales, there needs to be transparency and coordination.

  Horizontal systems integration in Industry 4.0 refers to the integration of the various software and hardware used in production. Regardless of how many different vendors the equipment and applications came from, it should work together seamlessly and intelligently. This enables Smart Factories to respond dynamically to new production requirements and helps facilitate predictive maintenance.

  Whereas Vertical systems integration refers to the integration of a manufacturer's various departments, from production to IT to quality assurance and sales. Rather than sailing data and making decisions based on limited data, each layer of the business can access all relevant information.

- **Cloud computing**

  Cloud Computing makes IT resources like data storage and computing power available on demand. It enables manufacturers to quickly scale their IT infrastructure up and down by paying for the resources they need, as per their requirement. This allows them to do things like create temporary test environments and keep IT resources local to customers when they deploy IOT devices as cloud infrastructure.

- **Additive manufacturing**

  Additive manufacturing is part of what enables manufacturers to rapidly produce small batches of products based on individual customer specifications. Using 3D printing and digital modeling, factories can profitably create one-off products. Since these products are created layer-by-layer, this process uses materials more efficiently, and the time to market is extraordinarily short.

Figure 2: Technologies driving Industry 4.0

https://doi.org/10.46335/IJIES.2024.9.4.7 e-ISSN: 2456-3463
Vol. 9, No. 4, 2024, PP. 35-42

3. BIG DATA ANALYTICS

3.1 Big Data

Big data is exactly what the name suggests, a "big" amount of data. Big Data means a data set that is large in terms of volume and is more complex. Because of the large volume and higher complexity of Big Data, traditional data processing software cannot handle it. Big Data simply means datasets containing a large amount of diverse data, both structured as well as unstructured.
The volume of data being produced every day is continuously increasing, with increasing digitization. More and more businesses are starting to shift from traditional data storage and analysis methods to cloud solutions. Companies are starting to realize the importance of data. All of these imply one thing, the future of Big Data looks promising! It will change the way businesses operate, and decisions are made. Following figure 3 shows various sources of big data [5].

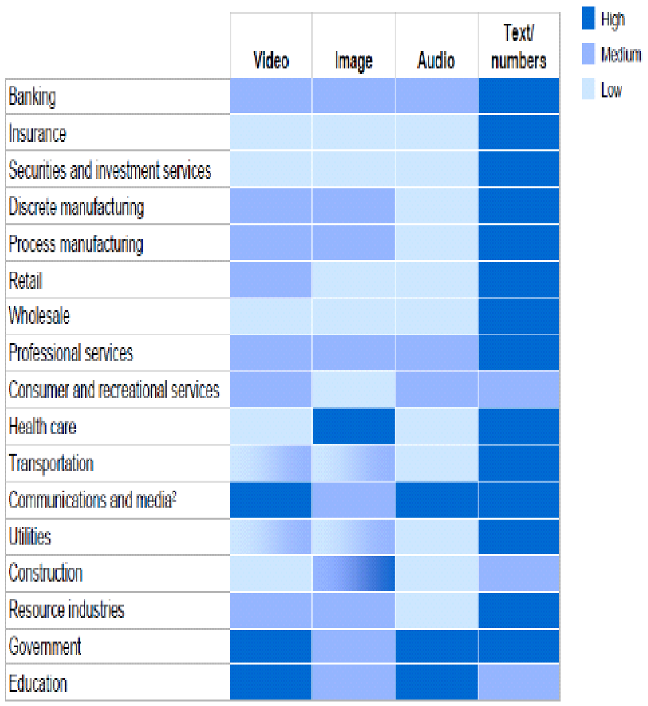

Figure 3: Sources of Big data

Big Data is described with the help of five characteristics known as 5v's. They are Volume, Variety, Velocity, Value and Veracity.

- **Volume** refers to the amount of data that is being collected. The data could be structured or unstructured.
- **Velocity** refers to the rate at which data is coming in.
- **Variety** refers to the different kinds of data (data types, formats, etc.) that is coming in for analysis.
- Over the last few years, 2 additional Vs of data have also emerged – value and veracity.
- **Value** refers to the usefulness of the collected data.
- **Veracity** refers to the quality of data that is coming in from different sources.

3.2 Big Data Analysis

This huge amount of data is merely useless if it not further use for any kind of process. Big data analytics allows organizations to use this huge data, process that data and draw meaning inference. Big Data allows companies to address issues they are facing in their business, and solve these problems effectively using Big Data Analytics. Companies try to identify patterns and draw insights from this sea of data so that it can be acted upon to solve the problem(s) at hand.
Big data analytics is the often complex process of examining big data to uncover information such as hidden patterns, correlations, market trends and customer preferences that can help organizations make informed business decisions.

On a broad scale, data analytics technologies and techniques give organizations a way to analyze data sets and gather new information. Business intelligence (BI) queries answer basic questions about business operations and performance.
Big data analytics is a form of advanced analytics, which involve complex applications with elements such as predictive models, statistical algorithms and what-if analysis powered by analytics systems.

Organizations can use big data analytics systems and software to make data-driven decisions that can improve business-related outcomes. The benefits may include more effective marketing, new revenue opportunities, customer personalization and improved operational efficiency. With an effective strategy, these benefits can provide competitive advantages over rivals.

3.3 Lifecycle Phases of Big Data Analytics (DBA)

There are total 9 phases as shown in figure 4, in the process of BDA[5]:-

- **Phase I Business Problem Definition –**

https://doi.org/10.46335/IJIES.2024.9.4.7 e-ISSN: 2456-3463
Vol. 9, No. 4, 2024, PP. 35-42

It learns about the business domain, which presents the motivation and goals for carrying out the analysis. here the problem is identified, and assumptions are made that how much potential gain a company will make after carrying out the analysis. Important activities in this step include framing the business problem as an analytics challenge. which helps the decision-makers understand the business resources that will be required to be utilized thereby determining the underlying budget required to carry out the project.

it can be determined, whether the problem identified, is a Big Data problem or not, based on the business requirements in the business case. To qualify as a big data problem, the business case should be directly related to one (or more) of the characteristics of volume, velocity, or variety.

- **Phase II Data Definition –**
  Once the business case is identified, now it's time to find the appropriate datasets to work with. In this stage, analysis is done to see what other companies have done for a similar case. Depending on the business case and the scope of analysis of the project being addressed, the sources of datasets can be either external or internal to the company. In the case of internal datasets, the datasets can include data collected from internal sources, such as feedback forms, from existing software, On the other hand, for external datasets, the list includes datasets from third-party providers.
- **Phase III Data Acquisition and filtration –**
  Once the source of data is identified, now it is time to gather the data from such sources. This kind of data is mostly unstructured. Then it is subjected to filtration, such as removal of irrelevant data, which is out of scope to the analysis objective. Here corrupt data means data that may have missing records, or the ones, which include incompatible data types.
  After filtration, a copy of the filtered data is stored and compressed, as it can be of use in the future, for some other analysis.
- **Phase IV Data Extraction –**
  Now the data is filtered, but there might be a possibility that some of the entries of the data might be incompatible, to rectify this issue, a separate phase is created, known as the data extraction phase. In this phase, the data, which don't match with the underlying scope of the analysis, are extracted and transformed in proper a form.
- **Phase V Data Munging –**
  As mentioned in phase III, the data is collected from various sources, which results in the data being unstructured. There might be a possibility, that the data might have constraints, that are unsuitable, which can lead to false results. Hence there is a need to clean and validate the data.
  It includes removing any invalid data and establishing complex validation rules. There are many ways to validate and clean the data. For example, a dataset might contain few rows, with null entries.
- **Phase VI Data Aggregation & Representation –**
  The data is cleansed and validates, against certain rules set by the organization. But the data might be spread across multiple datasets, and it is not advisable to work with multiple datasets. Hence, the datasets are joined together.
  For example: If there are two datasets, namely that of a Student Academic section and Student Personal Details section, then both can be joined together via common fields, i.e. roll number.
  This phase calls for intensive operation since the amount of data can be very large. Automation can be brought into consideration, so that these things are executed, without any human intervention.
- **Phase VII Exploratory Data Analysis –**
  Here comes the actual step, the analysis task. Depending on the nature of the big data problem, analysis is carried out. Data analysis can be classified as Confirmatory analysis and Exploratory analysis.
  In **confirmatory analysis**, the cause of a phenomenon is analyzed before. The assumption is called the hypothesis. The data is analyzed to approve or disapprove the hypothesis. This kind of analysis provides definitive answers to some specific questions and confirms whether an assumption was true or not.
  In **an exploratory analysis**, the data is explored to obtain information, why a phenomenon occurred. This type of analysis answers "why" a phenomenon occurred. This kind of analysis doesn't provide definitive, however it provides discovery of patterns.
- **Phase VIII Data Visualization –**
  A sort of representation is required to Obtains value or some conclusion from the analysis. Hence, various tools are used to visualize the data

https://doi.org/10.46335/IJIES.2024.9.4.7 e-ISSN: 2456-3463
Vol. 9, No. 4, 2024, PP. 35-42

in graphic form, which can easily be interpreted by business users. Visualization is said to influence the interpretation of the results. Moreover, it allows the users to discover answers to questions that are yet to be formulated.

- **Phase IX Utilization of analysis results –** The analysis is done, the results are visualized, now it's time for the business users to make decisions to utilize the results. The results can be used for optimization, to refine the business process. It can also be used as an input for the systems to enhance performance.

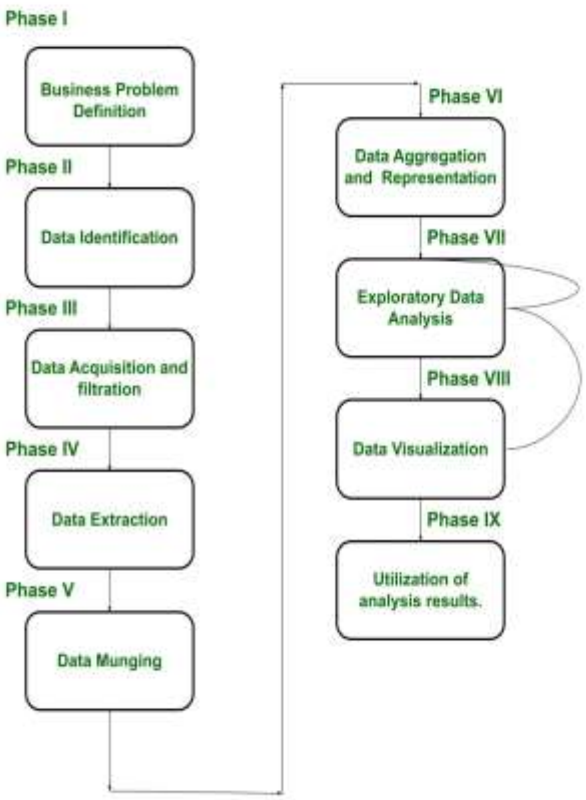

Figure 4: Life cycle diagram

It is evident from the block diagram that Phase VII, i.e. exploratory Data analysis, is modified successively until it is performed satisfactorily. Emphasis is put on error correction. Moreover, one can move back from Phase VIII to Phase VII, if a satisfactory result is not achieved. In this manner, it is ensured that the data is analyzed properly.

- **Big Data Analytics in Industry**

In the industry there is continuous generation of data through various industrial processes. Data is collected analyzes patterns are Detected. The model developed can be used to for testing new data generated. The process is repeated continuously as shown in figure 5 [6].

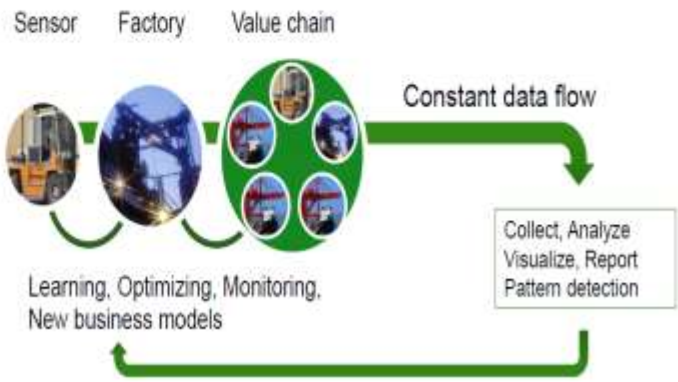

Figure 5: Big data analytics in industry

4.1 Big data analytics benefits in industry

The benefits of using big data analytics include:

- Quickly analyzing large amounts of data from different sources, in many different formats and types.
- Rapidly making better-informed decisions for effective strategizing, which can benefit and improve the supply chain, operations and other areas of strategic decision-making?
- Cost savings, which can result from new business process efficiencies and optimizations.
- A better understanding of customer needs, behavior and sentiment, which can lead to better marketing insights, as well as provide information for product development.
- Improved, better informed risk management strategies that draw from large sample sizes of data. Big Data analytics uses to identify fraudulent activities and discrepancies. The organization controls it to minimize the list of suspects or root causes of problems.

4.2 Big Data the core of the 4.0 industry

Industry 4.0 is the fusion of the real world with the virtual world. It combines the advantages of both technologies Big Data and Artificial Intelligence (AI). It combines both technologies to nurture automatic learning systems.

Manufacturers in today's marketplace seek to achieve business intelligence through the compilation,

https://doi.org/10.46335/IJIES.2024.9.4.7 **e-ISSN: 2456-3463**
Vol. 9, No. 4, 2024, PP. 35-42

International Journal of Innovations in Engineering and Science, www.ijies.net

analysis and sharing of data across all key functional domains in order to achieve production excellence. The industry make interconnection between systems and computers using IOT , now with big data analytics, it is able to analyze large amounts of data using intelligent machines that can make decisions without any human involvement [5][6].

4.3. Industry 4.0 be without Big Data

Modern factories are becoming increasingly complex and interconnected, creating new challenges that automation, driven by Big Data, can address. The large amounts of data resulting from the IOT and modern computer systems have enabled intelligent factories to quickly boost their efficiency and make significant gains in terms of increased uptime, accelerated production and reduced errors.

The possibilities of interconnection, facilitated by the IOT, have been creating very large and complex information networks. Without the Big Data, 4.0 technologies would not have been able to interpret and extract value from all that information in order to learn, generate predictive analysis patterns and operate autonomously and in such a precise manner. Therefore, without Big Data, there would have been no 4.0 industry, nor the intelligent technologies that support it.

4.4 Big data analytics challenges in industry 4.0

Despite the wide-reaching benefits that come with using big data analytics, its use also comes with challenges [6]:

- **Accessibility of data.** With larger amounts of data, storage and processing become more complicated. Big data should be stored and maintained properly to ensure it can be used by less experienced data scientists and analysts.
- **Data quality maintenance.** With high volumes of data coming in from a variety of sources and in different formats, data quality management for big data requires significant time, effort and resources to properly maintain it.
- **Data security.** The complexity of big data systems presents unique security challenges. Properly addressing security concerns within such a complicated big data ecosystem can be a complex undertaking.
- **Choosing the right tools.** Selecting from the vast array of big data analytics tools and platforms available on the market can be confusing, so organizations must know how to pick the best tool that aligns with users' needs and infrastructure.
- With a potential lack of internal analytics skills and the high cost of hiring experienced data scientists and engineers, some organizations are finding it hard to fill the gaps.

VI. CONCLUSION

In the current situation, the volume of data is growing along with world population growth and technology growth. This is a clear sign/indication of the increasingly widespread use and necessity of Big Data Analysis solutions. Big Data is not just a trend of technology, but it is a business practice that helps the industries/enterprise to stay in this competitive world to make their proactive data-driven business decisions to improve sales and marketing team performance, increase revenue.

REFERENCES

[1] *"An Introduction to Industry 4.0",*
[2] *https://www.emnify.com/blog/industry-4-0* .
[3] *"Industry 4.0 and the fourth industrial revolution explained", by Is coop*
[4] *"Industry4.0", http://www.rgbsi.com/industry4.0.pdf*
[5] *"India's Readiness for Industry 4.0",CII*
[6] *"Big Data Applications in Industry 4.0", P. Kaliraj and T. Devi*
[7] *"Big Data In Industry 4.0",*
[8] *https://nexusintegra.io/big-data-industry-4-0*