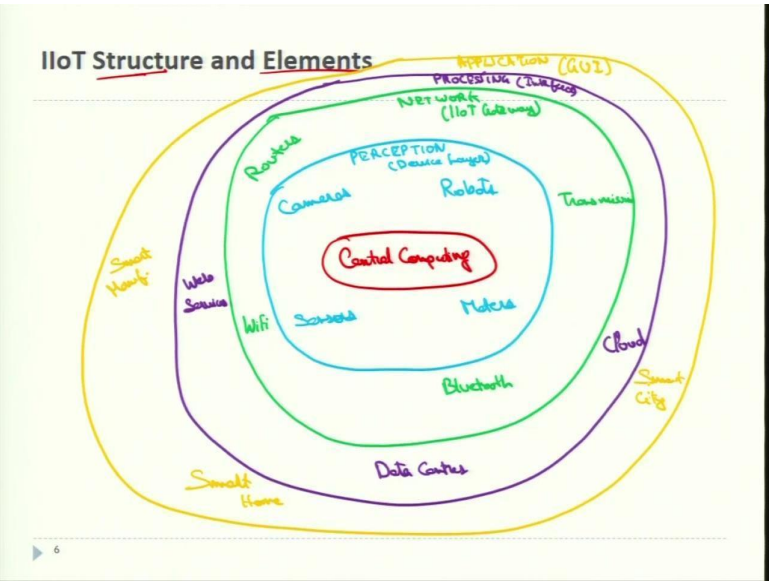
**Computer Aided Decision Systems - Industrial practices using Big Analytics**

**Professor Deepu Philip**
**Department of Industrial and Management Engineering**
**Indian Institute of Technology Kanpur**
**Professor Amandeep Singh**
**Imagineering Laboratory**
**Indian Institute of Technology Kanpur**

**Lecture 29**
**Structure of BDA in IIoT**

Welcome back to the talk on the Industrial Internet of Things. Whether we are talking about the wearables for health monitoring, or we talk about the tracking of the soldiers, tracking of the drones, that will go in the defense technology or so. The Industrial Internet of Things is all the rage nowadays.

(Refer Slide Time: 00:37)




So, there is a structure and elements of it. When I say structure, there are layers of the Industrial Internet of Things. If I tried to name those layers, they are perception, network, processing and application. So, I have first at the center of central computing.

Now, there are four layers as I said: perception, network, processing and application. Application is finally, when you try to monitor your health, when you try to monitor the health of your manufacturing concern. Health of the manufacturing concern means, your output throughput of the manufacturing concerns, time saving, or you are trying to monitor any other systems such as the GPS system or so. It is the outermost layer, the application layer that goes out, maybe a smart city, smart grid or so.

It is the ‘Central level’ or the ‘Perception layer’ that collects the data which are the actual sensor, the actual systems or the elements that try to collect the data. Here we have,

- Cameras,
- Robots, Sensors
- Meters,

We call it meter because the equipment’s used to measure the energy that is known as energy meter, so, what is the total energy, total power that is utilized for running this unit or maybe for running this computer, I can measure just putting equipment on the input wire, what is the energy that is being used. So, those are known as meters.

Sensors, they could be multiple sensors but tactile sensors, they could be proximity sensors, IR sensors, temperature sensors or so. Then robots can do multiple operations. Cameras can have visual inspection on visual or view or photographs with a system, so, this makes my perception layer. So, the perception layer is all the actuators, all the devices like I have mentioned here, cameras, robots, sensors, meters, which are used to gather the data. It is referred to as edge layer or device layer.

So, the Industrial Internet of Things, this is a very important layer where the controllers for the Internet of Things or the Industrial Internet of Things are also taking part in it and the process controllers are typically embedded and dedicated to specific functions.

For instance, a sensor, that is for example a temperature sensor, the function of that is to be able to record the temperature that would be sent to a web service or sent to a cloud through a second layer that we call a ‘Network layer’. The network layer we have,

- Routers-
- WiFi
- Bluetooth
- Transmission technology

Now, the network layer is the one which consists of dedicated IIoT devices that transmit the collected data over the Internet. So, following the flow of data, starting from the smart sensors and moving up, this next layer provides an IIoT gateway. This becomes my IIoT gateway, which is also known as Edge Gateway. So, which solver performs the local storing and processing of the collected data before sending it to the next level up in the network.

Edge gateways are not only meant for process control, but also they perform these operations or some of the calculating operations while sending it to the cloud as well. Sometimes small cleansing also happens. So, here the security measures are also implemented to shield or to secure downstream devices from any outside attacks as well. So, data centers which are there, where we try to collect the data to the next layer that is also taken care of to have a good security within this.

Next come in my 'Processing layer'. This is the processing layer. So, processing layer, we have,

- Web services
- Data centers
- Cloud

Now in the processing layers, we are in the main cloud domain now. At this point, the data has reached a database or a server in the data center like a zoo data center as they are underwater, as I mentioned, large amounts of computing power are usually needed to execute the data analytics or the algorithms which are there in machine learning. Therefore, the most critical data or data processing that happens in this layer really.

In addition to the regular algorithms that are making programs or developing a small interface between the next application layer or so. This layer also includes web services. So, in which all the cloud based technologies support this process, web services describe data transform from the device to the other interfaces of the programs. So, it could be the markup language that you are trying to develop. So, that gives you a final interface to the user. So, it is, I would say, an interface to the user device.

So, it is an important point to note here that in the processing layer, the data is stored in the cloud, and when it is stored in the cloud, the security is higher, because at the local level itself, it depends upon the security staff, the human intervention could also be there or may be the local computer is storing the data. So, storing it in the cloud or taking the data from the data centers through web services or so, that is happening in the processing layer. So, data centers are managed by a dedicated company and security is much more stringent than that is there in an local onsite system.

Next comes the last layer, where we have the final device that is the interface with the user, that is the 'Application layer'. Application layer refers to the Graphical User Interface (GUI),

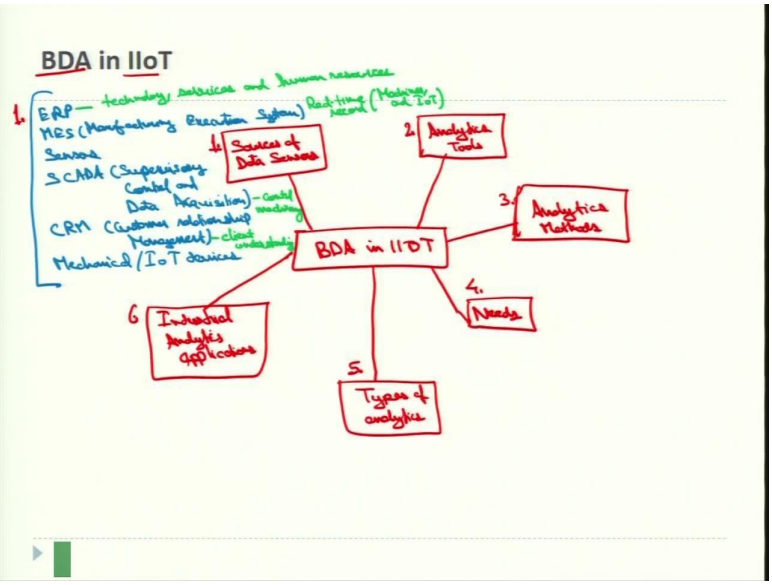
that is deployed, which helps the operators, the users or the humans to interpret the information, which has come as an outcome from the previous step.

So, in this layer, one can find user programs that generate reports, the analysis, the statistics, the plots, the graphics or so. There are many application layers in the market nowadays, but just to put or to name a few of them, it could be anything. We can have maybe,

- Smart manufacturing system
- Smart home systems
- Smart city

So, these are the major components of the Industrial Internet of Things. Here definitely you could see because so many things are connected. The data is to be transferred from one place to another, now this is structure and elements.

(Refer Slide Time: 09:57)




Next is, Big Data Analytics in the Industrial Internet of Things. The two domains, Big Data Analytics and IIoT, which are going to have a convergence. And despite the fact that IIoT and BDA each of them have been a subject of extensive research on their own. So, due to the high use of the sensing system and devices, it is widespread in industrial networks nowadays. Big data production in IIoT is obvious. So, complete integration and implementation of big data analytics processes in the IIoT systems are not yet possible to do concept of the IIoT systems still being developed, because we are talking about industry 4.0 and this system is so getting

matured that we even trying to reach to industry 5.0, where completely digital systems would be there, that is the minimal intervention of the human would be there.

So, already IIoT adoption, the integration of these systems into Edge cloud computing, the industrial marketplaces, big data virtualization technologies or technology advancements like the central processing systems, smart manufacturing, these are already there.

Here, I would like to discuss the structure of the BDA in IIoT. If I say BDA in IIoT, this can have various elements and sub elements in it. For instance, we can have,

1) Sources of the Data Senses
2) Analytics tool
3) Analytical methods are not only the tools, what kind of models would you like to choose.
4) Needs
5) Industrial analytics applications
6) Types of Analytics

1) So, in the sources of data,

- ERP is Enterprise Resource Planning,
- MES is Manufacturing Execution System,
- We have direct sensors here.
- We have a SCADA program that is the data system, which is Supervisory Control and Data Acquisition.

There are two kinds of data, supervisory and nonsupervisory. By supervised learning or supervised data we mean you train the machine using data that is well labeled, that is grade level tested, that is the data acts as a supervisor to the system or algorithm that is already tested. So, this is supervised learning or supervised base or supervised control data acquisition system.

So, in supervised learning the learning means training the data set by iteratively making predictions on the data and adjusting the correct answer. So, we correct the data, then we have

- CRM system, that is Customer Relationship Management.

So, these are just a few of many sources of data production in industrial settings.

- So, also here we could have mechanical or IIoT devices.

So, here ERP systems give businesses the ability to use a system that is made up of numerous integrated applications, which helps to manage business requirements and automating as well a variety of pack office tasks which involves technology services and human resources.

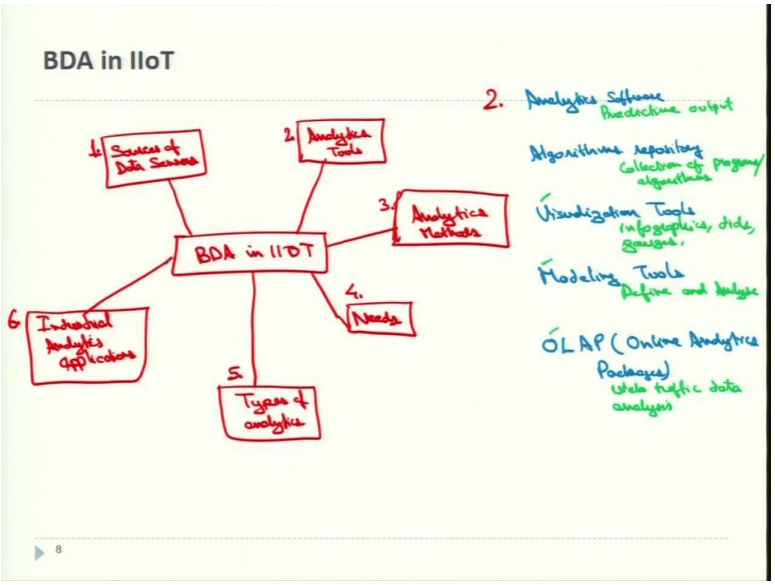
Next, the Manufacturing Execution System assists in maintaining a real time record of the manufacturing data and receives up to date information from robots, machines and IIoT devices. So, it is a real time record.

Next is the SCADA (Supervisory Control And Data Acquisition) system used in industries to monitor and manage a facility or maybe a piece of machinery. For example telecommunication, maybe water and waste control energy or in the pipeline, maybe oil or gas refinery or transportation. So, they try to control machinery. The MES systems are connected with machines and the Internet of Things.

Next CRM systems, CRM systems are frequently employed to manage the client relationships (Customer Relationship Management), that is industry are using these devices to carry specialized tasks which produce a massive amount of data every day.

So, by using analytics tools to the data gathered from different systems or machines or the devices that we try to use, user information can help in decision making and the data that is extracted from here could be used to have a good processing over it, where we use BDA and try to make a decision out of this. So, these act as a Decision Support System for the Internet or the Industrial Internet of Things.

(Refer Slide Time: 19:45)




2) Analytics Tools, which could be,

- Analytics software
- Algorithms repository
- Visualization tools
- Data modeling
- Online Analytics Packages (OLAP)

But to gain understanding from a sizable amount of the industrial data, a variety of analytics tools are needed. These tools include the modeling software, online packages, the visualization tools, the modeling tools or the software or the online analytics packages.

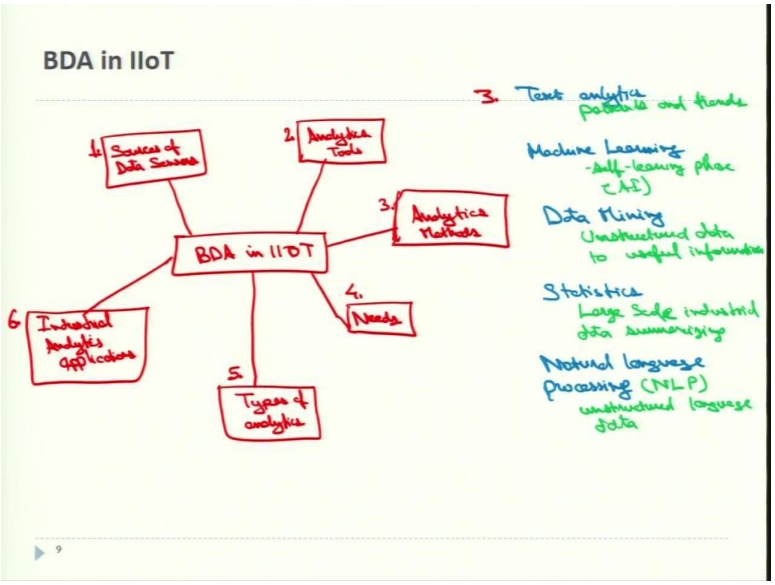
Analytics software helps us to have a predictive output that means it helps us to forecast unforeseen events.

An algorithm repository helps us to keep a record or collect the programs or the algorithms which were created by the data scientist or which were created by the analysts and which will guided by data engineer, which were supervised to some small extent and which were not used in past or which are sometimes tested in the past as well. So, this repository helps us to keep a record. I would say collection of programs or algorithms. So, which set of programming languages were used, which set of the standard user interfaces were there, those all help have a kept proper record of that and we can use them or retrieve them whenever required in the future.

The tools for visualization, these help in presenting data in different formats. For example, infographics or maybe the dials, the gauges which were developed or maybe the geographical graphs, the heat maps, the detail bar, the pie and fan charts, all this helps us to keep a record of the graphics or the plots which were made in the previous tests or previous runs. So, within the context of the corresponding information system in industries, the modeling tools play an important role which are used to define and analyze. The data requirements for supporting the business processes. So, web traffic data is tracked and analyzed using OLAP. These are the analytics tools.

However, they are also one of the subsystems of the big data analytics in IIoT as a complete system. So, these subsystems have sub subsystems and these sub subsystems do have elements in them, which are listed here.

(Refer Slide Time: 24:12)




3) Next comes Analytics methods. Analytical methods if I try to talk about, there are numerous methods which can help in extracting value from big industrial data, which enables quick and effective decision making. Now it could be ranging from,
4) Text analytics
5) Machine learning
6) Data mining
7) General statistics
8)
9) Natural Language Processing.

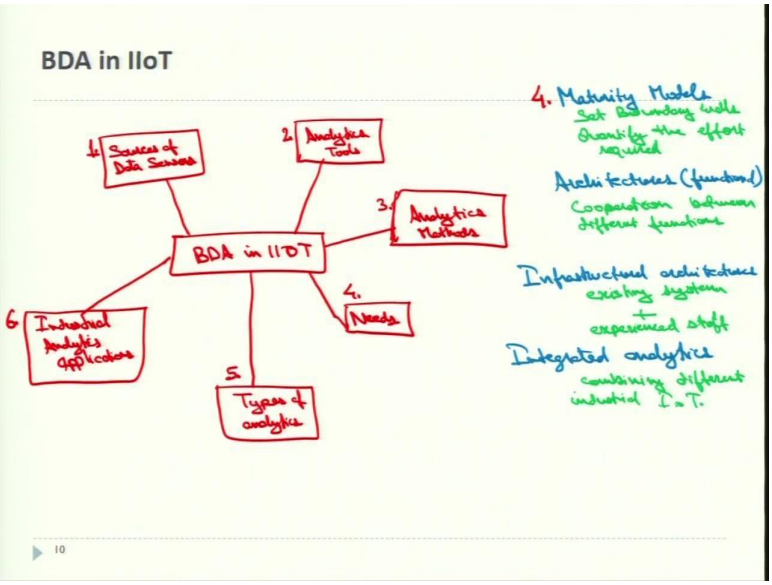
But through the use of statistical pattern learning, the text analytics reveal patterns and trends that helped to generate high quality information. This has been elaborated in the previous lectures as well. How the Internet of Things helps us to collect the data, that is, how do we try to have to track them if the programs are made to collect all the information regarding the specific product review that the person has given or the specific product multiple reviews of them people have given? So, what are the trends which are coming, what are the demands that people are expecting or of people who are returning to the Amazon requirement returning to Flipkart? So, why is the product being returned? They always ask the reason for that. They have a listed checklist. And from that checklist, they also asked the following questions: what could be the reasons, what do you expect in this product? Was it not fulfilling your basic requirement or did you change your mind or do you have a similar product at a lesser price? So, these patterns and trends could be tracked in text analytics, when we use analytics methods in connection to the sensors or the equipment or the machines which are there in my Industrial IoT.

Machine Learning means the systems can go into the self-learning phase without having a use to program them time and again , the AI systems are being developed. It helps us to have unstructured data so that useful information on large scale industrial data are gathered, summarized, analyzed and interpreted with the aid of the statistical tools, like I took an example of the data collected for the heat or the temperature, that is collected each second of the run of the engine. Do we need to have data for each second? Data could be collected in each millisecond as well, but do we need that, can we have an average of 1 minute.

If in that average what is the variation, we tend to have a standard deviation in that, that is a statistics. If from that average we try to say at what points the average is higher or lower, that also becomes more statistics. So, large scale industrial data could be summarized to be what is useful only using the statistics, but that is why identical methods always have statistics as one of the important subsystems.

Now, the NLP (Natural Language Processing processing) system. These are used in an industrial setting to extract unstructured data that is in the form of language.

(Refer Slide Time: 29:17)




4) Next comes the needs or the requirements that the system has. The needs or requirements we have with us are the first one as,

1. Maturity models
2. Architecture (functional)
3. Infrastructure architectures
4. Integrated analytics

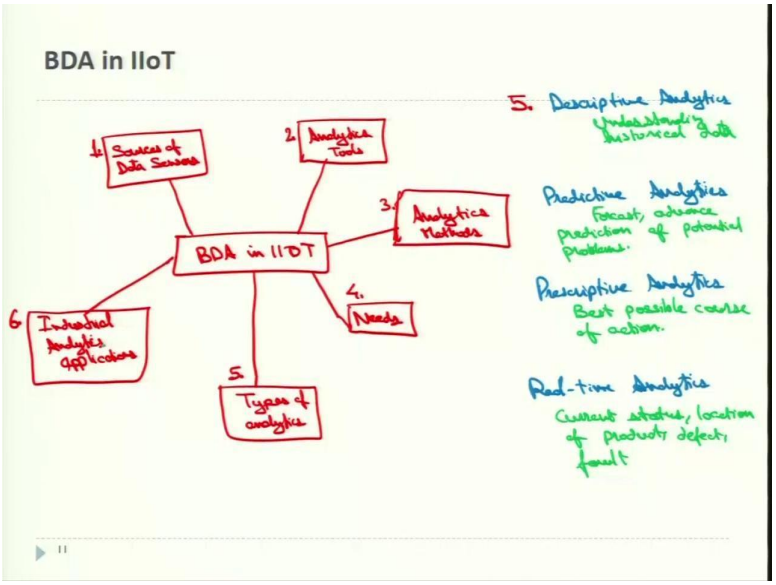
In these, what we have is, we have to create new analytic systems for the Industrial IoT where certain needs should be taken into consideration right at the time, right in the beginning. The model of maturity is one of them. Analytic system capabilities can be measured and tracked with aid of maturity models. So, these aid in quantifying the effort needed to complete a particular stage of the development. That means, we try to set boundaries as walls or conditions of where the model is going to be mature, then quantify the effort required. So, these models aid in keeping in organizations big data program in a good shape, that means, the system is planned, the model planning is done, where to close the model, what is the hypothesis, what size of the data we do we require, what were the size of this sandbox that we are developing analytics is there, that will be helpful.

Now, an architecture model, that is a functional architecture, aids in defining the rules and responsibility between the functions of analytic systems. That is how system components cooperate to carry out a specific system mission, which means analytic systems should be created to handle a massive amount of data in real time in an industrial setting. Cooperation between different functions is important.

Now, in this situation when the cooperation between different functions are required, infrastructure architectures also play a big role. Because designing a big data infrastructure from the pre-existing equipment using an industrial paradigm requires experienced scientists. So, because we have to use existing systems+experienced staff to handle them. Staff means the data scientists or data engineers or the big data analytics team, which are going to use the present infrastructure architecture only to develop a team out of it and to develop an analytic system out of it.

Now, supporting integrated analysis is also one of the needs where various types of Industrial IoT data is considered as one. So, it becomes the main requirement of a complete analytics system. I would say, combining different Industrial IoT systems.

(Refer Slide Time: 33:54)




5. Next comes, the Types of the Analytics. The types of the analytics we have,

- Descriptive analytics
- Predictive analytics
- Prescriptive analytics
- Real time analytics

So, understanding the historical data is the purpose of Descriptive Analytics. With the number of defective items in the past and the reason for the defects or so, this is all descriptive analytics.

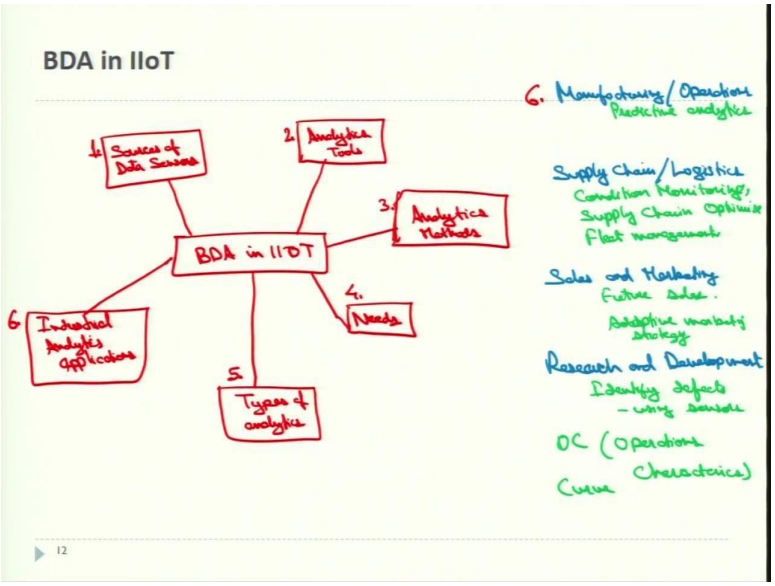
So, if I try to project that into the real time analytics, we could also mean to have maybe the current items. I would mention it first here, current status or current location of a product or a fault or a product, defect or fault.

Predictive Analytics is anyway the prediction that is what is going to happen in advance, that is forecast or advance prediction of potential problems.

Prescriptive Analytics last but not the least, advice or suggest the best course of action.So, that is whether the machine is receiving right material or not, whether the machine performance is consistent since the beginning or since the morning or not, whether in the morning itself the performance might be higher because of worker is fresh to work upon, in the evening hours whether performance is coming down, whether that is due to the human error because of workers are down or it is because of the machines had got heated up or between the change of the setups, do we have the pilot piece or the first piece that has been manufactured just after the setup change tested or not. So, do the output or do the data that is generated for the pilot piece same as that is being produced for the other pieces as well or not. So, this all is prescriptive analytics the best possible course of action. What could we do if the problem is here. So, this becomes my type of analytics, that is there. And kinds of the analytics also decide, what kind of the sensors or what kind of the data models are we going to use. So, this anyway is simple big data analytics, which we are not connected to different sensors and different sensors are helping us to collect the data in the different layers which I discussed.

So, in the network, in the processing, in the application, the data is being collected in different applications and finally in the applications whether we are trying to have a smart manufacturing system. It could be a completely transformed factory or maybe one site or one element of the factory is changed. So, if small changes happen, we can say yes, the color of the factory, in terms of the greenness, in the terms of the smartness was black first, it has not turned to complete green, but from the black it has turned to gray color or so. So, if the small portions will really change.

(Refer Slide Time: 38:47)




6. In Industrial Analytics Applications, we have different kinds of applications of the industry, that is,

- Manufacturing/operations
- Supply chain/logistics
- Sales and marketing
- Research and development

Now, these kinds of the different functions of the industry, whether in manufacturing supply chain, sales marketing, research and development are to be taken into account. Which function are we targeting, the machine performance history, in the manufacturing and operations, I have taken certain examples. For instance, the previous history of the machine, we can have predictive analytics here as well. So, whether rescheduling is required or maintenance is required or machine failure is there, that all becomes a part of a manufacturing or the operations of the system. So, in this Decision Support System, it can aid in industrial or controlled or manufacturing Decision Support System.

The logistics and supply chain can benefit greatly from effective application of analytics. For example, condition monitoring, then supply chain optimization, fleet management, statistics supplier management, this all becomes the kind of industrial applications that our Industrial IoT is having through big data analytics.

In sales and marketing, analytics tools help to allow the business to foresee or increase the future sales. For example, to help in determining the seasonal trends that can lead to developing an adaptive marketing strategy.

In research and development, analytics tools can help to identify the defective components, maybe through sensors, whether the components are defective or not, should it be shipped to the customer or not, even after being shipped to the customer, the test could be taken by the sales manager or after the sales manager there itself. So, we can identify defects using sensors. So, this helps us to develop a better OC curve, that is the Operations Characteristics curve.

This was a brief introduction or the broad discussion on big data analytics in the industrial IoT system. So, I would like to discuss case studies in the next lecture, where we will see how Industrial IoT is being supported by big data analytics to have an output, which industry has proven to be very helpful or very profitable to them. Thank you.