Machine Learning for Soil and Crop Management
Professor Somsubhra Chakraborty
Agriculture and Food Engineering Department
Indian Institute of Technology, Kharagpur
Week 1
Lecture 01
General Overview of ML and DL Applications in Agriculture

Welcome friends to this NPTEL Online Certification Course of Machine Learning for Soil and Crop Management. And today, we are going to start our first week of lectures and today will be our first lecture and in this lecture, we are going to talk about the general overview of Machine Learning and Deep Learning Application in Agriculture.

(Refer Slide Time: 00:52)

So, in this week 1, we are going to cover these following concepts. First of all, we are going to discuss about the Big Data and its storage aspects. Then, we are going to talk about Machine Learning, Artificial Intelligence in brief and Deep Learning and what are the interconnection between these 3 terms; Machine Learning, Artificial Intelligence and Deep Learning and then, we are also going to touch some aspects of Artificial Intelligence, their advantages and some disadvantages.

Also, we are going to talk about some traditional uses of Machine Learning, where we generally use Machine Learning? Some sectors where we use Machine Learning and then we will start the different types of Machine Learning approaches. We are going to talk about, we are going to basically give the overview of supervised method of Machine Learning and then we are going to talk about the unsupervised method of Machine Learning.
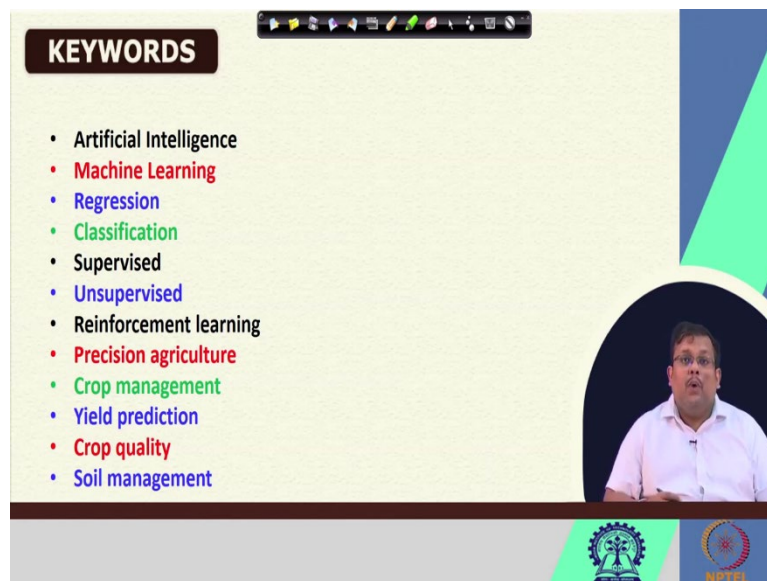
Then, we are going to talk about the semi-supervised areas of Machine Learning and also, we are going to discuss the reinforcement learning. Also, when we come when we complete these discussion or overview of these supervised, unsupervised, semi-supervised and reinforced learning, we are going to talk about the application why the application of Machine Learning in agriculture is, is a necessity, we are going to discuss that.

Then we are going to define what is precision agriculture? What are the different components of precision agriculture? And we are going to also talk about the Machine Learning for crop management, what are the different application of Machine Learning of crop management like crop yield prediction, then crop quality identification and then weed detection, so disease detection, we are going to talk about this, we are going to have a brief overview of this with some examples from different sources.

Also, we are going to talk about the Machine Learning, how Machine Learning can be used in soil management. In terms of use of some sensors for rapid analysis of soil properties, also, we are going to talk about how image processing and data or deep learning can be used for predicting different types of soil properties.

And also, we are going to talk about the sub hyperspectral remote sensing and how Machine Learning in combination of hyperspectral remote sensing can be used for different types of soil management practices. So, these are the major concepts which we are going to cover in our first lecture first week. So, remember that these will give you the basic overview. However, in subsequent weeks we are going to discuss them in details.
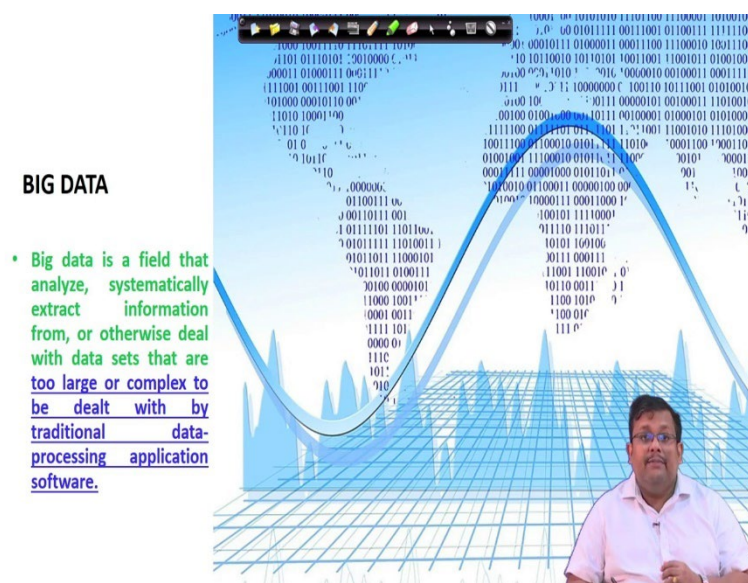
So, let us move to these are the some of the keywords which we are going to discuss, like Artificial Intelligence, Machine Learning, regression, classification, supervised, unsupervised, reinforced learning, then precision agriculture, crop management, yield prediction, crop quality, soil management in this week one of lectures.

Now, let us start with the first important description that is Big Data. Now, what is Big Data? Big Data is a field that analyse and systematically extract information from different sources and these data are too large or complex to be dealt with by traditional data processing

application software, so we require special storages or special application or special software to handle this data.

This data is huge data set and it requires huge amount of storage and it is not possible with the traditional methodology to handle this huge amount of data, we need some kind of help from the machine or help from several softwares for dealing with this kind of data. So, nowadays we are seeing the application of Big Data in different sectors. For example, in genomics, in genetics research, we have seen the application of Big Data for identifying or for explaining different gene related aspects.

And also, in case of soil also, you can see that nowadays, Big Data are used for creating the different types of management decisions, especially in digital soil mapping. So, we are going to discuss those later in this in this course, but remember, the Big Data is a huge and complex data, which cannot be dealt with traditional data processing application software.

(Refer Slide Time: 06:41)

So, Big Data has already I told you that it requires large volume, it has a large volume of data and which is produced by different types of digital technologies, this is important. Big Data is created by different types of digital technologies and it requires large storage capabilities in addition to editing analysing and interpreting. So, interpretation of the Big Data, it has a considerable potential, why we generally nowadays more rely towards Big Data?

So, interpretation if we interpret the Big Data correctly, that has huge potential to add value for society, environment and decision makers and different types of Machine Learning

application or deep learning application help us to analyse these Big Data for correct decision making in different applications.

(Refer Slide Time: 07:44)
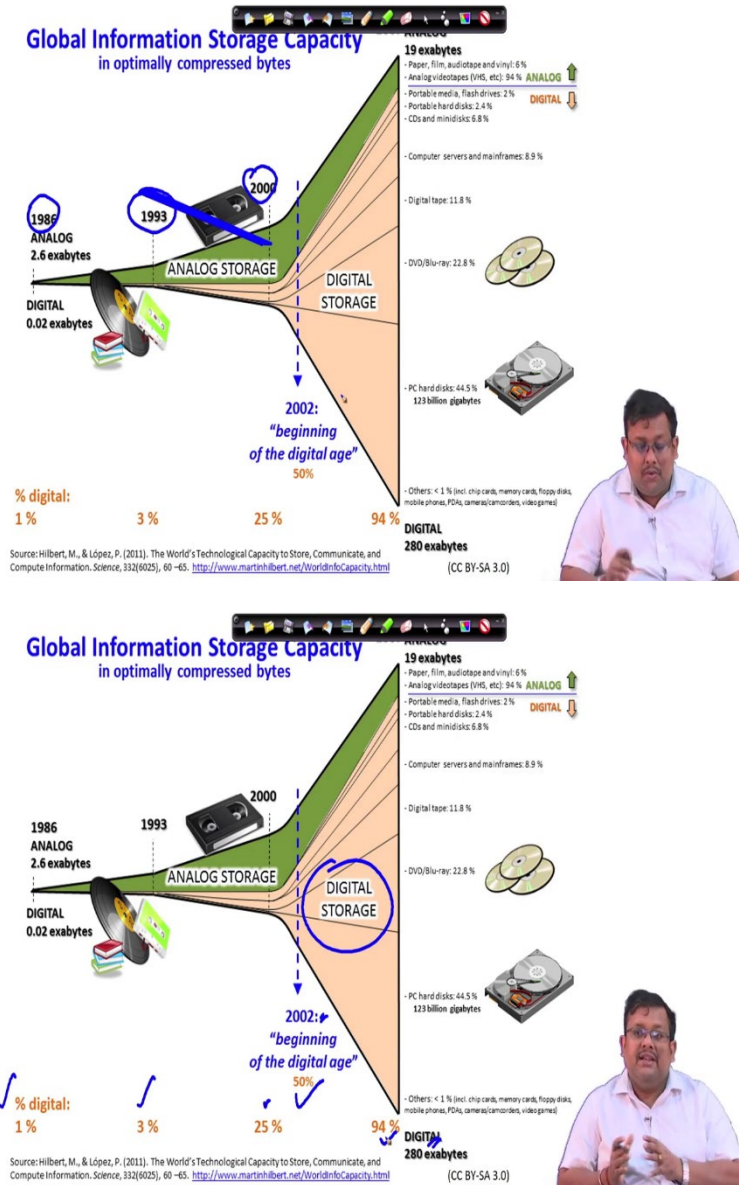
But, as I have told you, that nowadays this the storage of the data is an issue, because Big Data requires big volume of storage space. So, of course, the digital storage application, the importance of digital storage has becoming an important paramount issue. So, you can see this graph or this diagram shows a evolution of the analog and digital storage media or storage capacity starting from 1986 to 2007, it is a little bit old up to 2007.

So, you can see that during 1986 most of the storage system or storage capacity was analog and when starting from 1993 we have seen some emergence of digital storage and in 2000

starting from 2000 there has been a boom of use of different types of digital storage. So, boom of use of digital storage, and you can see that the use of analog storage continuously reduced after the year 2000 and in the year 2007.

We can see that 94 percent the storage capacity are being utilized by digital storage, whereas, only few amount, only 6 percent is being used by analog storage. So, the 2002 is considered as the beginning of the digital age from where we have seen a huge amount of increase in digital storage device. So, in where in 1986 we have seen only 1 percent digital storage.

In 1993 it went up to 3 percent and in 2000 it is 25 percent and in the year 2002 we have seen 50 percent of the data being stored in digital storage and in the year 2007 we have seen 94 percent of the data was stored in digital storage and nowadays in the year 2021 we have seen almost 100 percent of the data are stored in digital storage and very few amount of data is stored nowadays in paper, film and audiotape, vinyls devices.

So, we can see the temporal changes of, the temporal progress of digital storage and shifting from different types of analog video tapes or VHS tapes to the computer servers and then digital tape, DVDs, blu-rays, and hard discs and application of them. So, nowadays, Big Datas are being saved in different types of servers and also in Cloud Platform and also, we can save it in different types of hard disk.

So, these Big Data storage as more and more application of Big Data we have seen in different sectors, there has been a tremendous shift of different storage devices also, simultaneously we have seen.

So, what are the challenges of the Big Data? We have seen the advantages of Big Data, but what are the challenges of the Big Data, the challenges of the Big Data, on account of their so called 5-V requirements. What are the 5-V requirements? 5-V requirements is volume, we have already discussed it, then variety, different sources of Big Data, different digital sources, from where the Big Data is generated, the velocity, veracity or accuracy and finally, value.

So, these 5 are the major important aspects or challenges of Big Data and the conventional data processing technique are incapable of meeting or these 5 challenges of the Big Data and as a result, there has been emergence of Machine Learning application, because Machine Learning is the solution for these challenges of the Big Data.

(Refer Slide Time: 12:56)

So, now, before we discuss in details about the Machine Learning, let us have a brief overview of Artificial Intelligence also. What is artificial intelligence? What are some of the application of Artificial Intelligence? And then what are the advantages and disadvantages of AI we are going to discuss.

(Refer Slide Time: 13:25)

So, Artificial Intelligence is the theory and development of computer systems which are able to perform task normally requiring human intelligence, such as visual perception and then speech recognition, decision making and transition between the languages, when you use a

chatbot, that is an example of AI system. So, here you can see some examples. For example, as an web search, Google is an AI system also.

Also, some recommendation systems are examples like YouTube, Amazon, Netflix and also you can see that there are some human speech understanding for example, Siri and Alexa, you are quite familiar nowadays with these and finally, you know also if you consider the Tesla self-driving car, these are also an example of AI agent, which shows the AI applications.

(Refer Slide Time: 14:40)
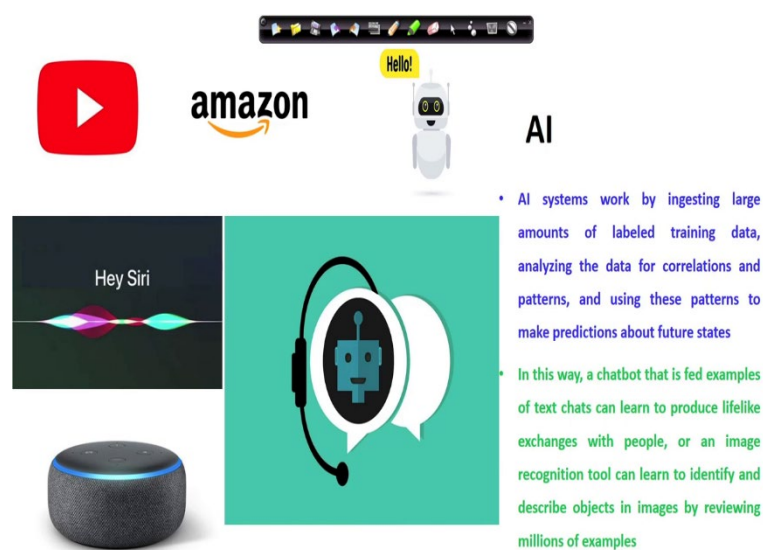
Now, so, how does the AI system works? So, AI system basically work by ingesting large amount of labelled training data, analysing the data for correlations and patterns using Machine Learning and using these patterns to make prediction about the future states. So, not only it accepts data from different sources and analyse it, it also takes decision resembling human intelligence that is why it is called artificial intelligence, to make predictions about the future states or also take intelligent decisions.

So, that is why it is called Artificial Intelligence and the whole thing is being modulated or operated by these artificial machine systems. So, that is why it is called the Artificial Intelligence. Now, in this way for example, when we talk about a chatbot, this way a chatbot that is fed examples of text chats can learn to produce lifelike exchanges with people, just like when we talk about each other.

When you feed these examples of textures in a chatbot then, it learns from there or an image recognition tool can learn to identify and describe objects in images by reviewing millions of examples. So, when there are millions and millions of images, this Artificial Intelligence system can scan them and identify the specific objects from them.

So, that is an example of what is Artificial Intelligence or identify specialist feature which we want to detect. So, it is called Artificial Intelligence.
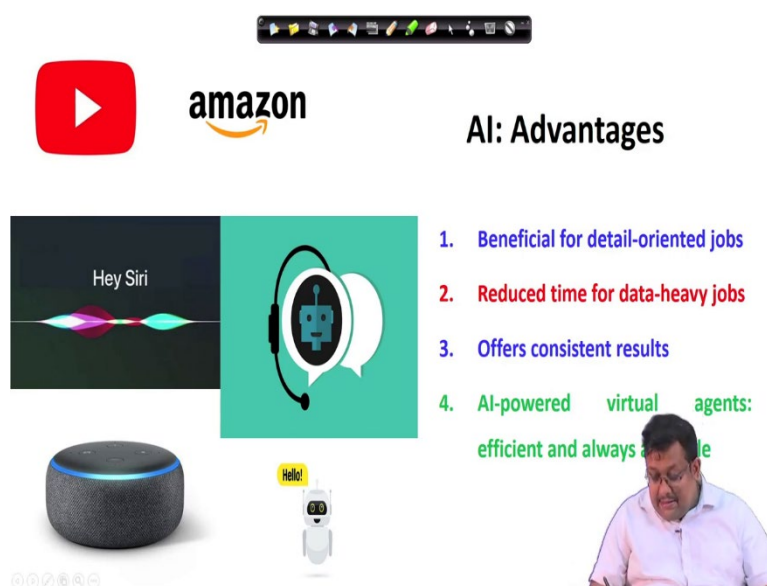
(Refer Slide Time: 16:36)
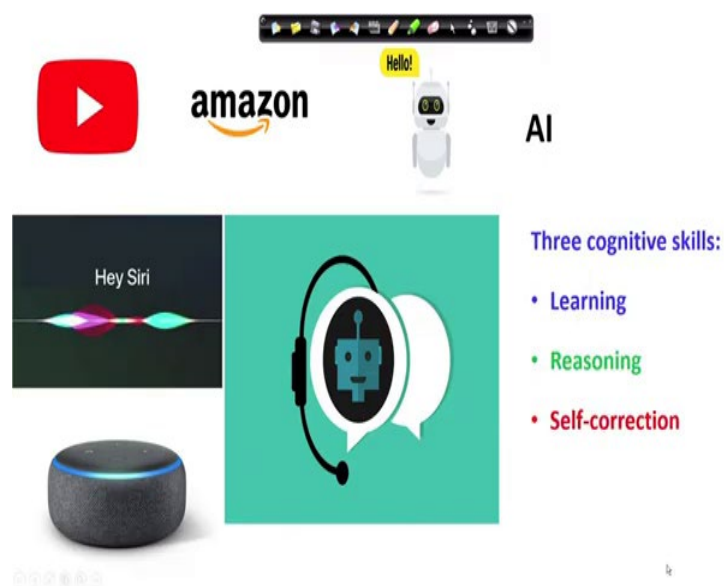
So, these terms Artificial Intelligence and Machine Learning are very much related and generally Artificial Intelligence depends on three cognitive skills. The first cognitive skill is learning, the second cognitive skill is reasoning and third word is self-correction. So, that is why not only learning which is which is which is generally a statistical learning or Machine Learning, but also it helps in reasoning and also it, it can self-correct itself.

So, that is why it is called Artificial Intelligence. Sometime people misunderstand that Machine Learning and Artificial Intelligence are same. No, there is a distinct difference. Artificial Intelligence uses some agents, whereas Machine Learning is more or less a statistical procedure.

So, an Artificial Intelligence has a more broader term, which can help not only, which signifies not only the learning of from Big Data, but also it can reason and also it can self-correct for future actions.

Now, there are different types of advantages and disadvantages of Artificial Intelligence. First of all, it is very beneficial for detail-oriented jobs. When you talk about some detail-oriented jobs, Artificial Intelligence is very-very effective. And of course, when we use Artificial Intelligence, it reduces the time of data heavy jobs.

We do not now have to do the data analysis using the traditional methodology of pen and paper, it can be automatically done by some softwares or by agents, Artificial Intelligence agents. So, that is why Artificial Intelligence can reduce the amount of stress for the data heavy jobs and it can handle that huge data using Machine Learning and also it offers consistent result this is another very good thing.

The consistent result is important in several sectors, where a small amount of mistake can have greater implication. So, Artificial Intelligence can help in getting the consistent result. Finally, Artificial Intelligence power virtual agents can, are they are very much efficient and they are always available. So, these are the four major advantages when we talk about the artificial intelligence.

(Refer Slide Time: 19:18)

Now, apart from advantages the Artificial Intelligence it has some inherent disadvantages also, what are these disadvantages? First of all, it is expensive or costly and secondly, it requires deep technical knowledge to handle or to build a AI agent. So, it has, we have very, still now we have very limited experts to build the AI based tools.
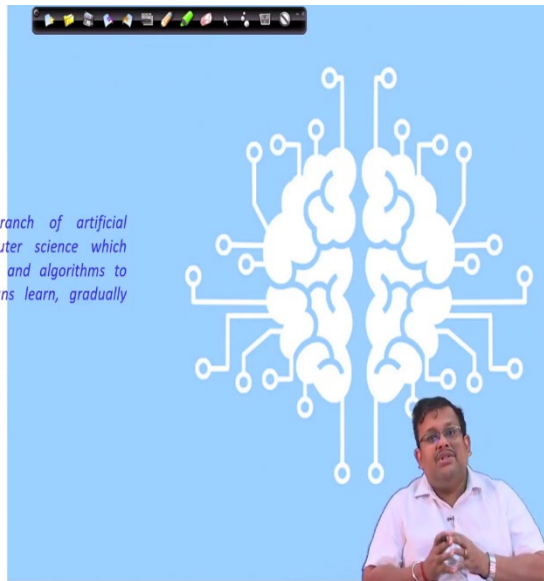
It is very specialized, and it required a deep technical knowledge to interpret the results or to analyse using artificial intelligence. And finally, it lacked the generalization from one test to another, this is another major drawback of the artificial intelligence. So, we have covered what is Big Data? What are the storage issues? And what are the evolution of storage for last 30 to 40 years.

And then we have learned about what is Artificial Intelligence? What are their applications? What are their advantages and disadvantages? And what is the relationship between Artificial Intelligence and Machine Learning.

(Refer Slide Time: 20:35)

Now, let us focus on Machine Learning now, since it is the major focus of this course. Now, Machine Learning is a branch of Artificial Intelligence. Remember, I am again focusing that Artificial Intelligence is a broader te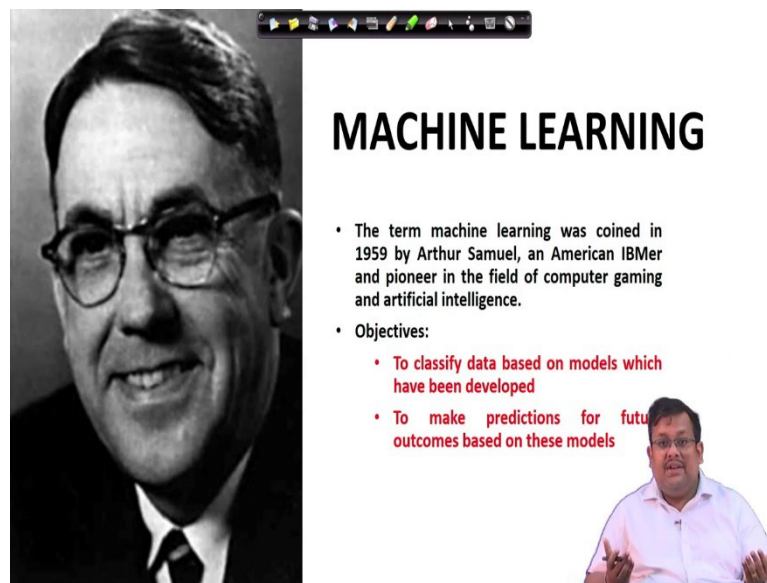rm; a subset of Artificial Intelligence is Machine Learning. Now, Machine Learning is a branch of Artificial Intelligence and computer science, which focuses on the use of data and algorithms to imitate the way that human learn gradually improving its accuracy.

Again, these Machine Learning learns using statistical methods just like the human learns from the data and gradually it improves its accuracy. Remember one thing that Machine Learning is also known as sometimes statistical learning, because most of the procedure which is employed Machine Learning or statistical procedure.

Now, who coined this term Machine Learning? This term Machine Learning was coined in 1959 by Arthur Samuel, an American IBM worker and pioneer in the field of computer gaming and Artificial Intelligence. So, he first coined this term Machine Learning and we will talk about the objectives of Machine Learning. What are the objectives of Machine Learning?

The objectives of Machine Learning are mainly 2, first of all, to classify the data based on models, which have been developed and secondly, to make the prediction for future outcomes based on these models. So, these are the two major objectives for Machine Learning. First, classification of the data based on models which they have developed and prediction. So, Machine Learning can be broadly classified into prediction and classification. Let us see the difference later.

(Refer Slide Time: 22:55)
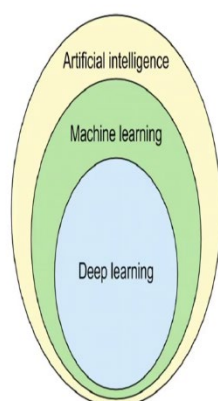
Now, again, if we want to know the interrelationship between Deep Learning, Machine Learning and Artificial Intelligence, then this figure gives you a better understanding. Machine Learning learns and predicts based on passive observations, whereas, artificial intelligence implies an agent interacting with the environment to learn and take action that maximizes its chance of successfully achieving its goal. So, this is the difference.
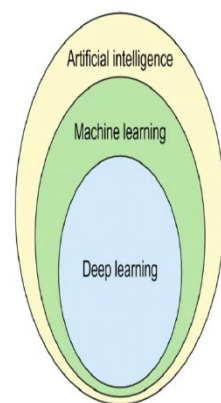
I have already told you that Artificial Intelligence is always agent base and it interacts with the environment with the help of Machine Learning to learn and take actions that maximizes its chance of successfully achieving its goal in future, because it can self-correct itself. So, you can see that Machine Learning is a subset of Artificial Intelligence and deep learning is a further subset of Machine Learning.

Now, what is deep learning, deep learning for example, nowadays we also use a buzzword called artificial neural network, convolution neural network, recurrent neural network. So, these are all deep learning methods. So, what is a deep learning, which is further a subset of Machine Learning.
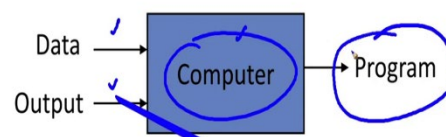
So, it is a part of a broader family of Machine Learning methods based on artificial neural network with representation learning. And this learning can be supervised, semi supervised or unsupervised. So, what we have learned that deep learning is a subset of Machine Learning. And Machine Learning is a subset of our Artificial Intelligence.
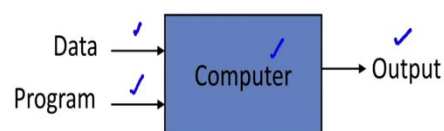
So, if we want to contrast the traditional programming and Machine Learning using a very simple difference, this gives you the basic understanding. So, we can see here that in case of traditional programming and both the Machine Learning, I mean, both traditional programming and Machine Learning we require computers. Now, in case of traditional programming, we are inputting the data as well as the program and the computer gives us the output. So, this is the traditional programming, but what about the Machine Learning?

In case of Machine Learning, we input data also we input output. So, apart from data or input, we also incorporate the output and it selects the most suitable program. So, this is the difference between a traditional programming and Machine Learning. So, in Machine Learning, it is not a simple programming traditional programming, where we already, we incorporate the data as well as a program which you want to run and it gives us the output based on that program.

However, in case of Machine Learning, that Machine Learning involves inputting both data as well as output, the output could be either labelled or unlabelled and it selects the program for analysing the data itself. So, this is the difference between traditional programming and Machine Learning. So, guys in this first lecture, we have so, far discussed some broader terms, we have tried to discuss briefly what is artificial intelligence?

What is Machine Learning? What is deep learning in broadly and we try to see their interrelationship and we have seen the difference between traditional programming and Machine Learning. So, let us wrap up our first lecture here. In the next lecture will be more, we will be talking about different types of Machine Learning like supervised learning and also what is regression? What is classification?

What are the supervised learning what are unsupervised learning and so on so forth. So, I hope that you have learned something new in this lecture. Let us meet in our next lecture to discuss different types of Machine Learning approaches in details. Thank you very much.