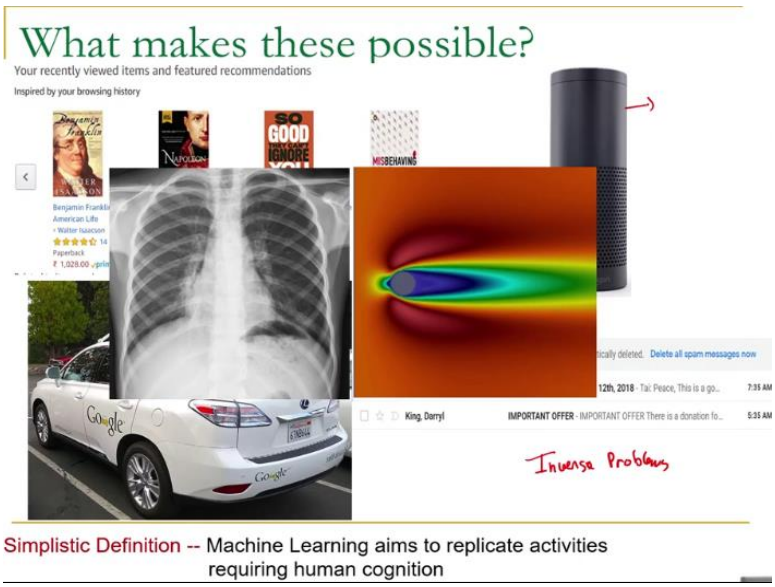
Inverse Methods in Heat Transfer
Prof. Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology – Madras

Lecture - 44
Overview of AI and ML

Welcome back, in the previous video we had seen why is it that inverse methods or machine learning is relevant to something like inverse methods. in this video I would like to give you a brief overview of these common terms that you have heard artificial intelligence and machine learning etc. and by the hope by the end of this I hope you can see why it is that machine learning itself is a type of inverse technique.

(Refer Slide Time: 00:48)



So, these are some common examples that I typically give whenever I talk about artificial intelligence or machine learning and these are somewhat older examples actually just over the last few weeks the field of AI and the field specifically of machine learning has really accelerated. so, we have a lot more examples that show practical strong practical results within the field but let us talk about these which should be common to everybody in amazon if you go you buy a book you get some book recommendations. there is amazon echo which recognizes your voice or you have Siri on your iPhone in case you have an iPhone or you have Alexa all these other sorts of things which recognize voices. then you have your spam detection within your email, you have self-driving cars a simple definition of ai or machine learning is you try to replicate activities, where usually human cognition human thought human recognition was required and you can apply this to radiologists which professor Ganapati Krishnamurthy of

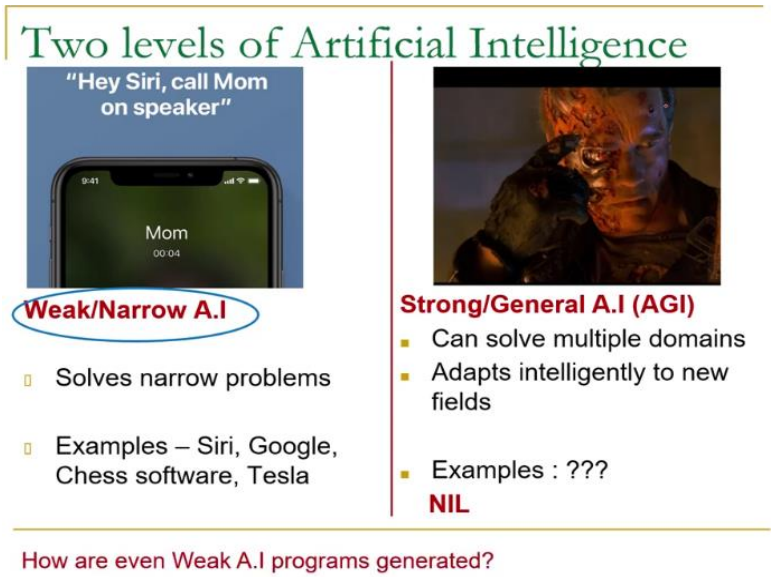
engineering design and other faculty members and even I a little bit We all are trying to do same techniques can you apply it for medical image diagnosis or my own specialty try to apply it to computational fluid dynamics and heat transfer.

Now what happens is all these things seem like entirely different problems, but the power of mathematics is such that all these different problems can be put within one single framework, that that is the remarkable thing these different problems can be set on one sim a single framework. This is somewhat analogous to how for a long-time people thought that you would need if you if you read early books that you will need different techniques for let us say booking your ticket that will be a different AI you know doing calculations on a computer that is a calculator okay you are going to call somebody that is a phone.

But now everything seems simply sits on one machine. why does it sit because digitally you are able to represent all these different types of information just as a bunch of zeros and ones. so just as digitizing unites various representations like images sounds everything is done by a single machine. similarly, machine learning unites different algorithms. so the algorithm for driving a car can in a strange way be united with the algorithm for recognizing your voice and the magic here is something that I have already discussed effectively. we can be proud of the fact that in fact you know already most of the moving pieces for this simply by treating these as inverse problems. so, if you see a sound speaker or a sound recognizer, you can look at it this way.

If I speak something, can you amplify it okay yeah that is a straightforward problem. But doing the inverse problem if you hear the amplified voice of somebody can you figure out what they were saying in words, that is the inverse problem. so, this as it turns out all these are special cases of inverse problems as I hope to convince you by the end of this week.

(Refer Slide Time: 04:07)

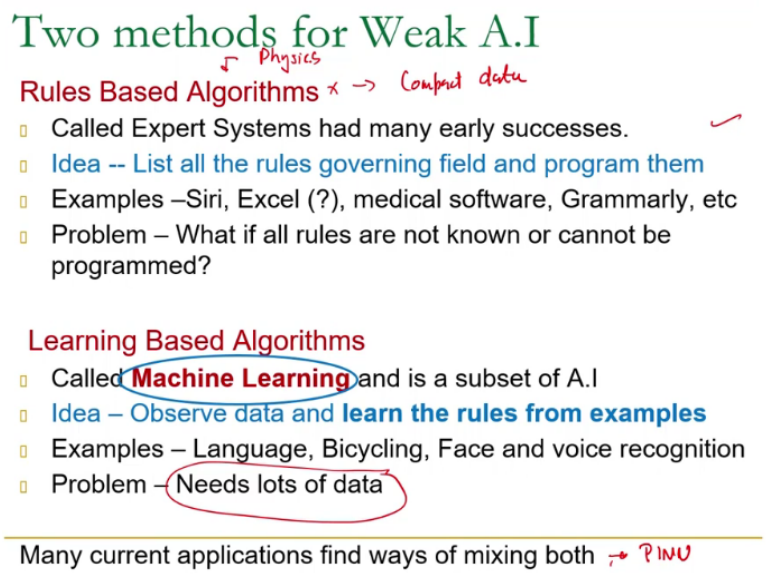


So, we have the large term AI and AI has these two splits or two levels. one can think of I am going to casually use these words, they are not the way I am using it might not be technically acceptable to everybody strongly within the AI field but this is good enough for a common level of knowledge. so weak or narrow Ai solves very specific problems.

So, for example, one can think of though today people might disagree, some type of Siri or even a simple google search in the earlier days pre ml days has a simple narrow problem. so if you ask Siri to do something or your phone to do something it can do only one thing if you ask it to do something else either it has to search the web but it can never do something complex like suddenly your voice recognition system will not become a chess playing system. okay it will not suddenly become something that will do medical diagnosis, whereas a human being on the other hand represents what is a theoretical sort of people, this is like the black hole in heat transfer. The strong or general AI is something that is able to do everything that a general human being can that is, it can solve for multiple domains.

It can adapt intelligently to new fields that a child who's barely speaking today can ride a cycle tomorrow can play chess and can watch video games or play video games etc. so obviously there are no examples of this other than in movies okay. so, we are going to concentrate on weak or narrow AI.

(Refer Slide Time: 05:51)



And if you look at weak ai historically, there are two types of algorithms that you have. you have rules-based algorithms these were called expert systems. the idea is very simple you list all the rules of governing the field and program them this is strictly speaking even this is not necessarily just an expert system on an ai there are other things going on of course this is vast field I do not want to simplify it beyond what it is.

But for our initial understanding there are things which you can very specifically talk about in terms of rules. for example, a calculator okay, it has finite set of rules, even playing chess it has a finite set of rules. you can actually encapsulate it a finite set of rules not for playing well but for simply playing chess. so, if you look at excel or Grammarly or earlier versions of grammar-based software earlier versions of medical software all these are rules-based algorithms okay. but what happens if you do not know what all the rules are or they are too complex to be programmed. for example, what are the rules governing walking, what are the rules governing cycling, there is no simple set of rules that you can give what are the rules in fact governing English grammar and what are the rules governing what something whether a mail is spam or not.

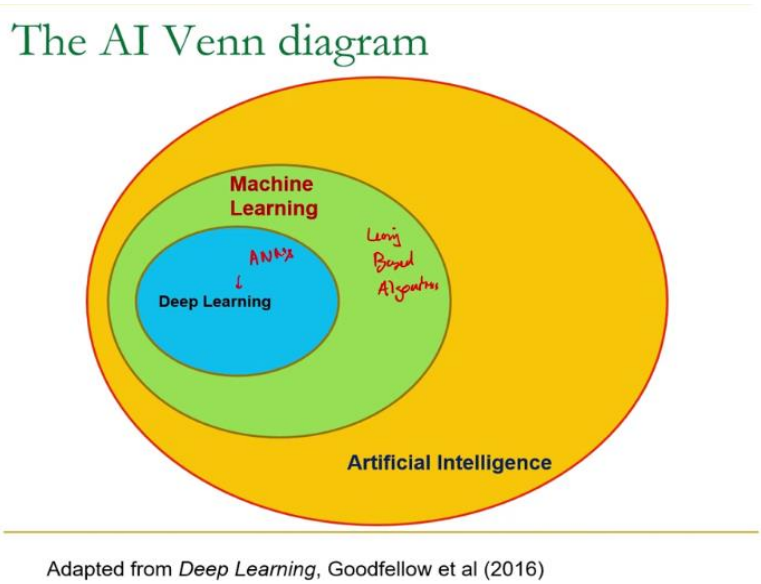
All these things cannot be encapsulated in a simple finite set of rules that is when people switch to learning based algorithms. so, this is basically a subset of ai. the idea is here instead of giving the rules you make the machine or the algorithm learn the rules from through examples. so, for example language how does a child or how do you learn language, how do you know what is grammar, you are not told in fact when we try to learn language as adults we are told if you

learn let us say a new language. let us say I try to learn Punjabi or I try to learn a foreign language like Spanish okay so when I try to do that usually people will give you know specific rules.

This is the way it will go there he sees; she sees, they see. so, stuff like that when you when you say you give a finite set of rules but you will see that will not be sufficient to cover all the possible examples. however, children learn language very rapidly in fact you and I speak the language of our parents fairly fluently without ever having learned explicit in a rules-based fashion so bicycling face recognition. how do we actually recognize? how what one person's face is or how do we recognize the voice of our mother all these are learning based algorithms, if you do not give rules there is only one problem this needs a lot of data. so, either you give rules or you give data now this problem we will come back to when we come to heat transfer.

Either you give me the physics if you do not give me the physics the data you give me better be really comprehensive. It has to cover a lot of ground because I am trying to infer the rules okay so the advantage of rules is this can be compact in terms of data. the advantage of learning is it can do fairly complex cases, but you have to give it a lot of data of course many applications generally find a way of mixing both and we will come back to this when we come to physics informed neural networks towards the end of the course. we are going to do a similar approach there as well. so, we will concentrate on machine learning. we are not going to concentrate on rules-based algorithms in some sense the physics-based algorithms we looked at can be thought of as a rules-based approach. now we are going to go to a learning-based approach

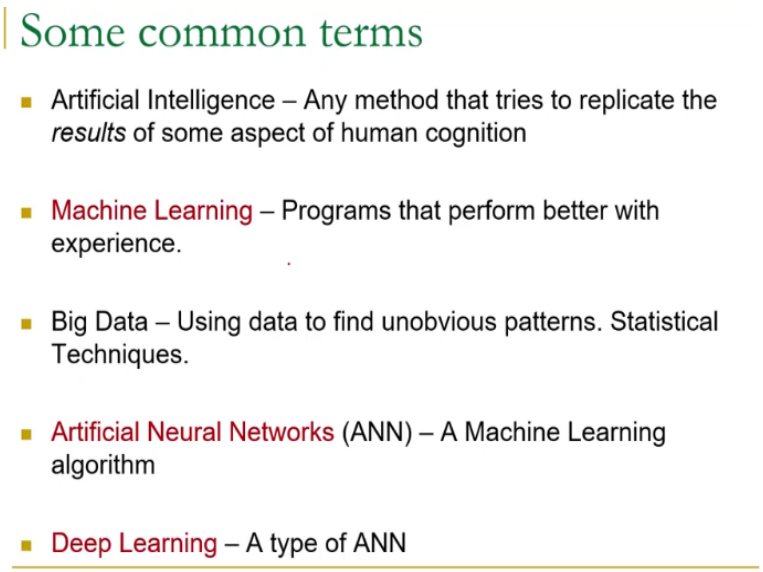
(Refer Slide Time: 09:49)



so overall this is what is known as the AI Venn diagram AI is a broad field info you know I am not covering all the other non-rules-based algorithms here obviously I am not even going to be covering learning-based algorithms very well here. it is just a very short time that we have. but within that a small area is machine learning which has learning based algorithms and we are not even going to cover all learning-based algorithms we are going to look at a subset which is deep learning or which I am going to call artificial neural networks which is also now not quite accurate a term the terminology keeps on changing within this field.

So deep learning is a specific subset of machine learning, so even though I said machine learning, we are actually going to be looking at deep learning very specifically.

(Refer Slide Time: 10:41)

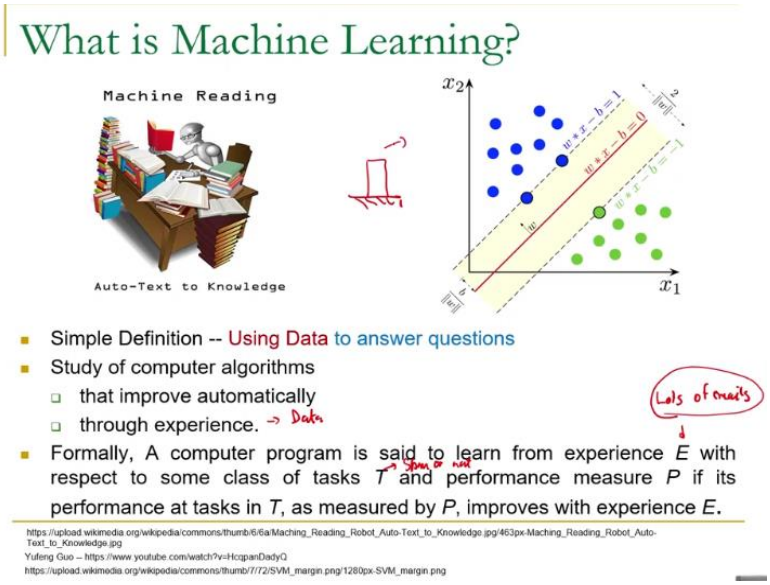


So here is some common terms AI is any method that tries to replicate the results of some human cognition, whether you want a machine to start seeing or machine to start recognizing people or you wanted to recognize spam from not spam or whether you want to recognize start walking or do some robotics all that is in some set, sense a set of subsets of AI algorithms. machine learning specifically is the set of programs that perform better with experience that have some learning involved in them. you would have heard the term big data. Big data is typically statistical techniques.

Big data has now get gotten subsumed within machine learning but that is not necessarily. so big data is a good fine independent field and it is actually mathematically much stronger than just ml. then we have ANNs they are a type of subset of a subset of machine learning algorithms.

these are parts of part of deep learning algorithms, they are a specific machine learning algorithm.

(Refer Slide Time: 11:47)



Now what is machine learning okay I have been talking about this. obviously, it is not a robot reading and doing some things. it is using data to answer questions now this should again trigger the memory that we this is precisely what we were doing with inverse methods also. okay so it is a statistical technique in in one sense of the term.

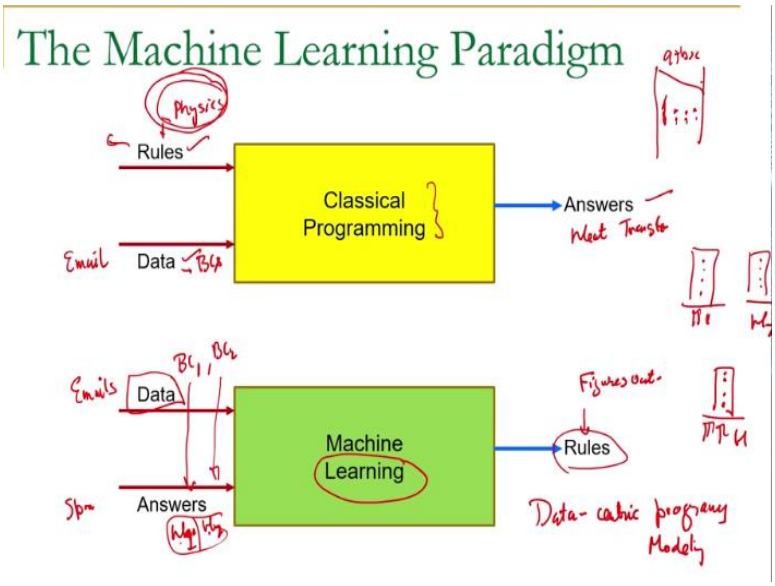
It is the study of computer algorithms that improve automatically through experience or basically through data. so, I am going to use a non-heat transfer example here just to ground you in general machine learning and when you come back to heat transfer, you will see how well these two actually combine very naturally. so generally, a computer program is set to learn from experience. so let us take a common example. let us say you are doing some spam classification. so, you get experiences lots of emails okay and the task is finding out whether it is spam or not. so, what we expect is if I have a spam classifier which is good, then its performance metric how many spams it classifies correctly, how many correct emails it lets through all that will increase as I show it more and more emails.

Same thing with any human being if you have an engineer and he is supposed to do some tasks well. let us say a person and a software engineer supposed to code well the more and more you give this person coding problems, this person's performance supposed to increase. this is the formal definition of when learning happens. now compare this to a standard cfd software, if I show it lets say flow past a cylinder or I show it a fin and ask it to find out heat transfer, it will

give me a certain heat transfer and the next time I ask it the same problem, it will take the same amount of time to give me the same answer, it can neither give me a refined answer is it say last time I gave it the answer I gave you the answer to ten percent accuracy, this time I can give it within five percent or within one percent or last time I gave ten percent accuracy in one hour, I can do it in thirty minutes now,

so, there is no learning that takes place within a typical cfd or software or typical heat transfer processes whereas what happens with ML is, if you give it more data, if you give it more data in the interim it is supposed to give you better and better answers.

(Refer Slide Time: 14:29)



So here is what is known as the machine learning paradigm a typical programming. so classical programming is like this, I give you the rules, I give you the data you are supposed to give me answers. so, for example I will again go back to the email case suppose, I give you an email and I tell you OKAY if it says, it is offering money it must be spam, so rules.

So whatever rules you want to build. it is selling medical bills, it is spam. so, these are explicit rules. but you know that not all emails that give you medical bills or offer you money it could actually genuinely be coming from your bank or some FD deposit in your name whatever it is okay so you cannot make these simple rules and give answers. but nonetheless if you are doing some classical programming that is what you concentrate on. if a program does not work well. so let us say you have a heat transfer problem the data here is whatever boundary conditions you have and the rules here are your physics of the problem and the answers you want let us say you want the heat transfer.

Let us say you write a Program and the answer goes wrong where will you intervene you will go back here and say did, I model it correctly was my code correct maybe I put a laminar model, when it should have been a turbulent model that is where you will intervene now compare this with machine learning okay. so, machine learning is like this. you give data so I will give emails I will give answers that is I will say okay here is email one this was spam email two this was not spam you will see this within your spam filter within Gmail or any email service that you use. that it will in case you have a new mail you can actually have the option of saying okay you put it in my spam folder but this is not spam. so, when you do that there is learning that happens.

How learning happens is Ironically or conveniently for us exactly through an inverse process, which we will see shortly. okay so you give an email you give answers spam, not spam, spam, not spam, you keep on giving it this and after labeling a lot of it figures out the rules. okay now what would be a heat transfer equivalent? the heat transfer equivalent will be I will give boundary condition one and I will tell you heat transfer heat transfer one I will give you boundary condition two I will give you heat transfer two. now you are supposed to figure out what is it that was happening in the middle that mapped pc one to heat transfer one pc two to heat transfer two. so, in this problem I could give it one set of measurements and say okay heat transfer was so much another set of measurements and say what the heat transfer or something else.

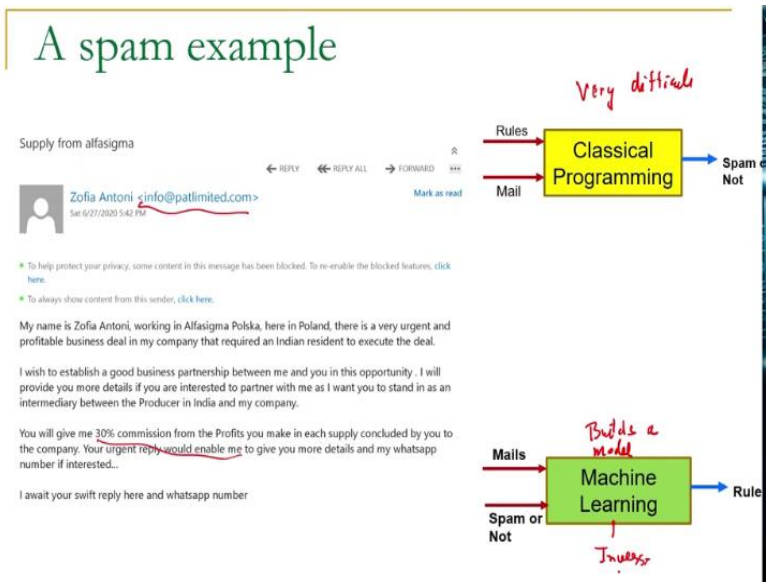
So finally, when you give me a new set of measurements, I should be able to say what the heat transfer is I hope this portion is clear so that is where learning differs from classical programming now one important thing. the important thing is in in case of the classical programming case in case your cfd software does not work you will go and try to fix the model. you will say okay maybe I put the physics wrong. now in case your machine learning model does not work one of the typical places to intervene is you can fix this, because this was the output. you have to fix the data okay either you labeled it wrong either you gave this heat transfer wrong or you have to start fixing the data you can of course fix the model too.

But this is what is known as data centric programming or data centric modeling more precisely in our case. we try to see did I give it sufficient data okay maybe I gave it too few examples of what happened or maybe my data was corrupted in some way. so, it is a slight change in

perspective of how we look at things. so, suppose I gave I found the conductivity in this case and it turned out to be completely wrong, then you have to wonder did I take enough thermocouple measurements or was my thermocouple too noisy. this is the kind of question you will ask rather than start questioning, whether your model a plus + b x is wrong that is not where you will put the question. usually you will/shall put the questions on maybe I made incorrect measurements, maybe my data was corrupted.

So that is a slight change in perspective between classical programming versus a learning-based approach classical programming concentrates on the rules. A learning-based approach concentrates on the data and its correctness and its statistical soundness etc.

(Refer Slide Time: 19:20)

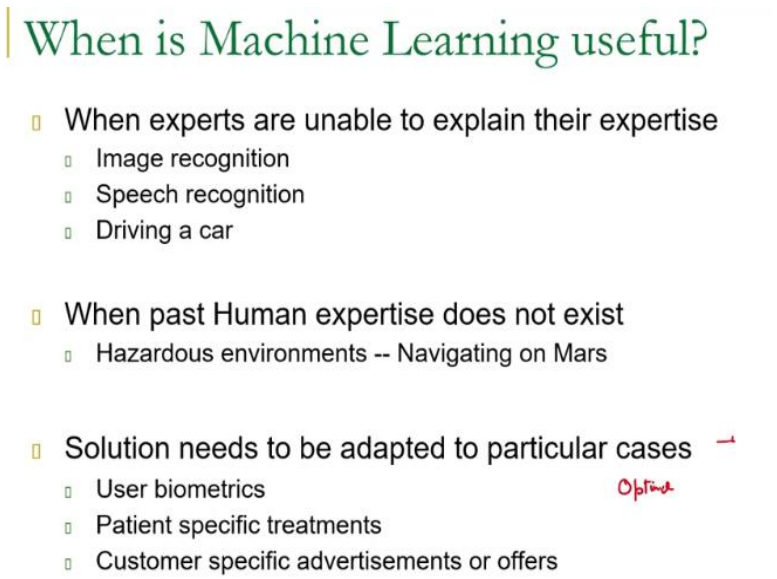


So let me give a simple example which I gave earlier based on spam. so, this is an email that I received a couple of years ago and this is spam so I am feeling free to just share it. you can see its grammar is good there is nothing clear in this which label this as spam. the problem is it says it is giving me some thirty percent commission and just by looking at it and knowing at the fact that I do not know this person the email address is fine. I know that this is spam.

Now it is very difficult for a classical programming thing to classify this, very difficult, you cannot give sufficient rules to find this out whereas for a machine learning based approach it would have seen millions and billions of such emails sent to lots of people and based on that it builds a model. okay it builds a model within itself that says okay this is actually spam and this is really how our spam filters work. so, as I show you the process it will become more and more you will start wondering how is it that the simple process it turns out I am going to put an inverse

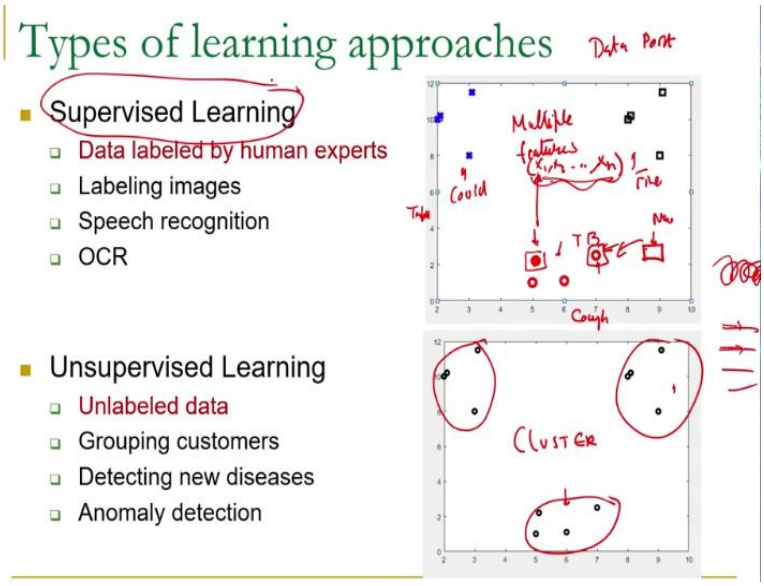
model here and that simple inverse model using a neural network tends to work fairly well which is kind of a miracle.

(Refer Slide Time: 20:37)



Now when is ml useful ml is useful typically when experts are unable to say, why their expertise works. okay so for example you look at an image and say cat or dog or person how is it that you are doing it is kind of hard to explain. okay or you need to adapt solutions to specific cases so this is something that is of great engineering interest. can you have data centers, where based on local climate, based on local geometry of the room, based on what the server is placed like, can you have optimal let us say ac position which can even be adaptive. so, in medical fields like user biometrics patient specific treatments all these are things that are slowly taking off there is still some time to go but this is the sort of future that you can imagine, you can use machine learning techniques to have very specific solutions for very specific adaptive problems rather than one solution for everybody even though it does not adapt well to them.

(Refer Slide Time: 21:42)



So, there are two types of learning approaches we are going to concentrate on what is known as supervised learning and that is really what we have been doing throughout in inverse problems. so, what we call data labeled by human experts. so I would like you to slightly step out of your comfort zone and think about this case where we have emails, we have labeled them as spam or not spam you know a machine does not know it by itself, that this is spam or not spam or we can think of each data point as let us say a covid patient or a person that has come to a doctor and let us say on the horizontal axis You have something like cough frequency, on the vertical axis let us say you have the temperature of the person and there are three types of patterns that we see low cough, low temperature, high cough, high temperature or low cough, high temperature, something of that sort you could see multiple patterns between people

And the doctors have labeled this as let us say this person has covid, this person has let say tb, and this person is fine okay obviously for the data that I have given that would not work out it is just some random image a thing that I have made up so you go through you go to past medical history give it to doctors and they label all this data and a new person comes. so, this person is a new person and now you do not have to go to a doctor you simply measure this person's cough frequency and temperature and say it looks like this person is a tb patient because kind of it looks close to that data.

so, this would be an example of supervised learning notice something that I have done an entire patient has been reduced to a single data point. okay so there is a trick involved here, which I will talk about in the next video, there is a trick involved here. so, but the same thing will happen for images too we will represent images just as a single point in multiple dimensions. so, in this

case for one patient I just measured two things cough and temperature I could have measured other things this person, height this person, weight this person, you know maybe what their image looks like of their lungs etc., What was their RTPCR test.

So just like what we did with regression, this could have multiple features. that is even though I am showing this as a single point in two dimensions, obviously there are multiple things that define one single patient. so, this could person could have x one x two a whole bunch of features that define this in this case we have taken two. similarly, an image has multiple pixels. so typically let us say we take a very small sixty cross sixty image then you have three thousand six hundred pixels which I will come to in the next video, you have three thousand six hundred pixels and you have three thousand six hundred pieces of information all that puts this person as a single point. now why all this?

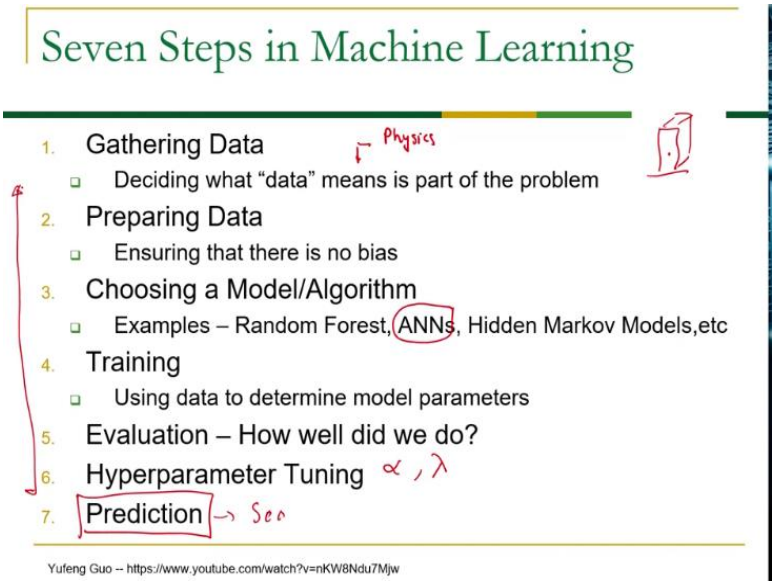
All this to say that supervised learning is mapping an input of this sort to an output in this case you are mapping it to tb covid or person is completely healthy. but that label is actually given by the doctor so the data is actually labeled by human experts so labeling images speech recognition in fact all the problems that we have done so far of regression. all those are examples of supervised learning. on the other hand, you could have unsupervised learning in this case you have unlabeled data. it is the same set of data you collected peoples just you collected peoples cough frequency and their temperature, but there was no doctor to label it you are let us say in a village and we do not have sufficient people this is an important problem that exists that we do not have sufficient data or we do not have sufficient people to label medical images

Because radiologist expertise is expensive in terms of their time, they are not free enough to label these things so let us say you have something of this sort then what do you do even though you cannot call this person tb this person covid and this person healthy you do know that somehow these three people are different. you can create a cluster automatically this would be like let us say we collect heat transfer data or we collect velocity data, pressure data from laminar flow and turbulent flow even a person who does not know turbulence or who does not know fluid mechanics or heat transfer can look at that velocity field and say hey this velocity field looks very different from the type of velocity field that goes straight.

So this is an automatic clustering the question is can we write algorithms that sort of mimic human intuition in such cases and actually go beyond human intuition. so, in this case the data

is not labeled but you can start grouping so in typical machine learning applications they can group customers they can detect new diseases, let us say you want to detect new patterns a cluster has broken out for covid somewhere or this person is showing atypical behavior anomaly detection some fraud detection all these cases are cases where we use unsupervised learning. but again, we can use this within heat transfer too I am not going to discuss that within this course. we are going to discuss primarily just supervised learning techniques.

(Refer Slide Time: 27:21)

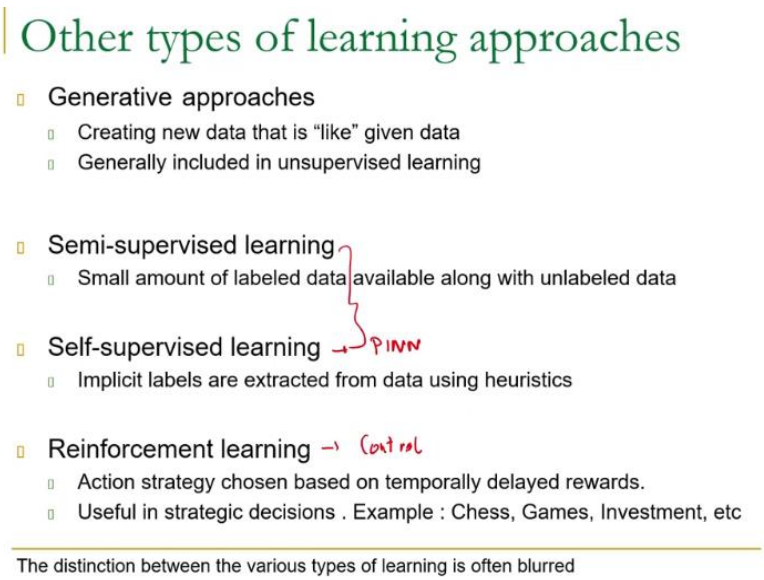


So, there are seven steps typically in a machine learning process. first of course you have to gather copious data. what is data what is not data when I simply say fin you should know enough physics to know that what I need to connect or the dimensions on the of the fin, I need to know where I have placed it what the time is, even here the physics of the problem actually comes in. what you collect and what you do not collect as relevant or irrelevant data. then you prepare data ensure it is clean like I said thermocouples could be noisy actually detailing the variance of each thermocouple. then you choose a model or an algorithm. so, there are various machine learning algorithms finally we are going to go with ANNs.

we also looked at Linear model's quadratic models etc. then you do what is known as training. okay so this training actually determines what the parameters are and finally you evaluate what the parameters are and we have already seen this idea of a hyperparameter when we had learning rate when we had the regularization. so, you need to tune these in order to work well for this problem. finally, here is where you make a prediction. this is the thermal conductivity of the material, this mail is email is spam or not spam, you actually use it like we are using it on a phone use an iPad or something or that sort it automatically recognizes your face. it is actually

making a prediction at that point. so, the important step of course this is the step that we see but these are all the hidden Steps that go towards making the algorithm.

(Refer Slide Time: 29:09)



Apart from the two supervised and unsupervised algorithms that I have talked about, there are multiple other types of learning approaches. there are generative approaches which are becoming very popular and I expect these to start affecting fluid mechanics heat transfer very soon in fact probably within the year. generative approaches are when you see three images of a cat rather than just telling me that the fourth image is a cat can you draw a cat. so that would be called a generative model. after having seen three flow fields can you generate a new flow field which is just like that would be the example of a generative approach. semi supervised learning just some amount of data is labeled.

Go back to the previous slide to understand that you can think about it. in some cases, I am telling that this person has covid and other people I was not able to actually get a doctor's opinion, but I still have their data the data is not useless. self-supervised learning is you actually boot strap I will not get into it but pinn effectively is an example of both semis supervised as well as self-supervised learning. as I have mentioned here the distinction between these various types of learning is often very blurry. the final thing is something called reinforcement learning a very powerful technique in order to control. remember I had talked about controlling temperature. so how do you control things how do you get a strategy for control is part of how reinforcement learning works.

So, in the next video I will cover what is known as the learning paradigm. I talked a lot about machine learning doing this or that, but how exactly does it work. so that is what I call the learning paradigm and you will see when you come to the end of it that actually the learning paradigm is exactly identical to an inverse problem. so, I will see you in the next video thank you.