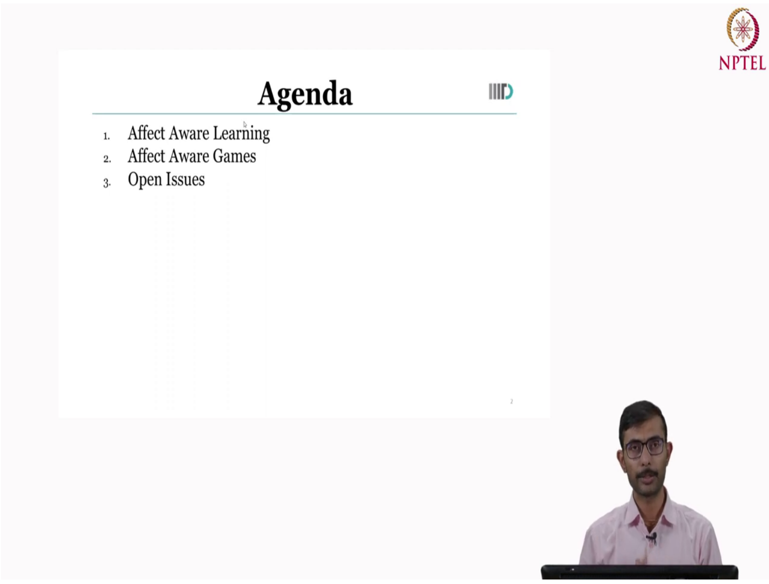
Affective Computing

Prof. Jainendra Shukla
Department of Computer Science and Engineering
Indraprastha Institute of Information Technology, Delhi

Week - 10
Lecture - 01
Part - 01
Emotionally Intelligent Machines: Challenges and Opportunities

Hi friends. So, welcome to this week's module. In this week, we are going to talk about Emotionally Intelligent Machines, Challenges and the Opportunities. So, so far, we have already discussed a lot about how to process the emotions, how to elicit the emotions and what are some of the applications that can be there.

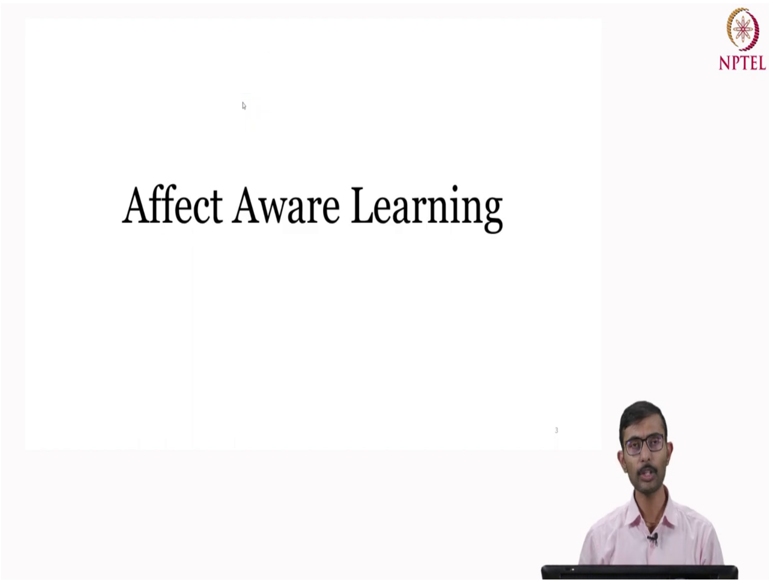
(Refer Slide Time: 00:45)



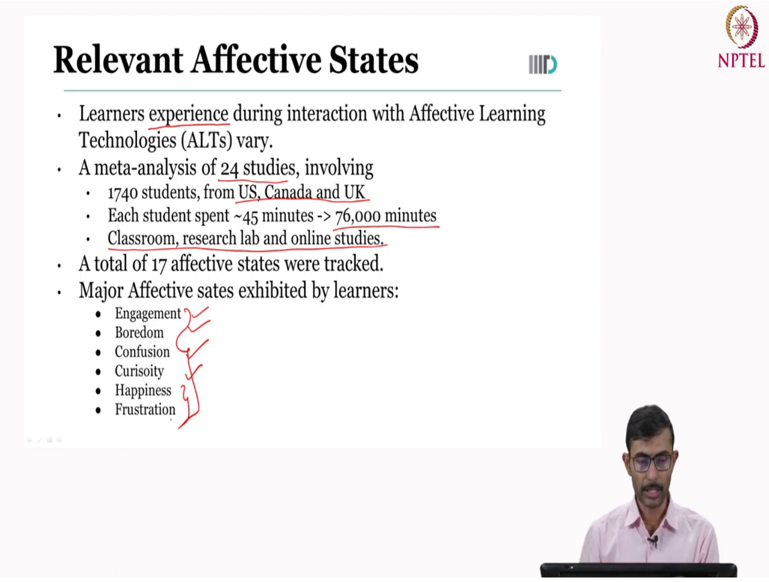
Today, we are going to take a step further and we would like to understand that about certain domains in which these can be used and what are certain open issues. So, 1st, we will be discussing the affect-aware learning. 2nd, we will be discussing the use of affective computing in games.

And 3rd, we will discuss in this week about the open issues that are there, and some indications on how to solve them. Particularly, our focus would be on the online recognition and the adaptation of the emotions while discussing these fields. With that let us dive in.

(Refer Slide Time: 01:27)



(Refer Slide Time: 01:31)



So, first affect-aware learning that is the use of affective computing in the emotional agents or in the emotional settings, in the emotional educational settings. So, it has been widely used in the learning technologies, in developing the learning technologies and it is also known as the affective learning technologies, where the emotion aware learning happens, those type of tools are known as the affective learning technologies.

And it has been seen that the learners experience during the interaction with these type of technologies it varies a lot. When we say the learner's experience, it means not only their emotions, but also their overall learning affects. So, please pay attention that overall, what we would like to do, we would like to sorry overall our aim when we are interacting with the learning technologies is to improve the learning.

So, then when we talk about the experience of the learners, it is not only the type of the emotions that the learner's experience, but also what is the improvement or the gain in the learning while interacting with these type of technologies. Otherwise, these affective learning technologies they will not be very fruitful.

So, first we would like to understand that what are the different types of emotions that can occur during the interaction with the affective learning technologies and accordingly maybe we need to develop the technologies. So, for this meta-analysis of 24 studies was done by some researchers which involved 1740 students from different countries including US, Canada and UK.

In these meta-analysis, each student roughly spent around 45 minutes and hence we had a total all together data of around 76,000 minutes which is significant data that we have. And these studies were collecting the data, not only from the classroom settings also in the typical research lab settings and also during the online settings. So, then we have a wide range of application settings in which these learning technologies were employed and the studies were conducted.

Now, it was found out that a total of 17 different types of affective states were tracked during these studies. Now, of course, 17 is a big number, and we would like to understand that we were the major affective states. So, if you look at the major affective states, then there were these 6 affective states that were exhibit exhibited by the learners and hence who are found to be most useful during the interaction with the affective learning technologies. So, of course, it is not a surprise engagement and boredom.

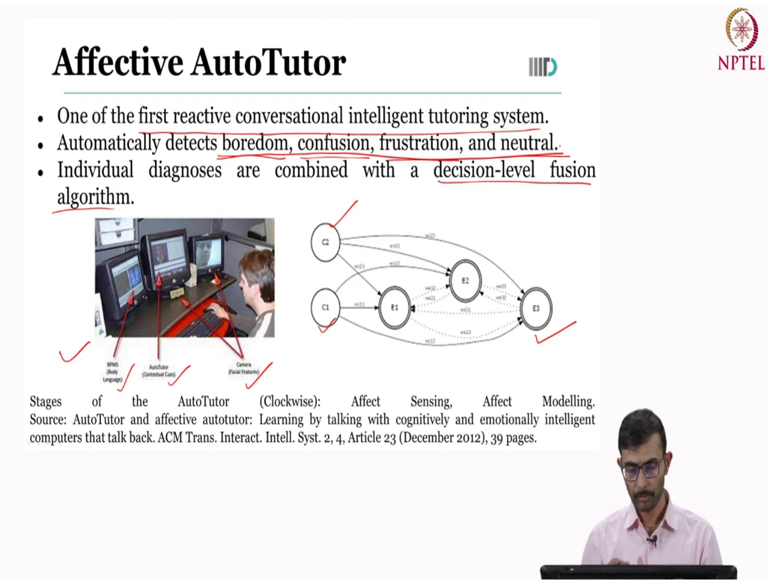
So, basically, engagement and boredom of course, you would like to understand and monitor whether the student or the individual interacting with this learning technologies is engaged or is bored. Of course, confused whether there is a complex topic that is being taught and hence it is making the individual the student confused about it. Whether the individual is more curious to learn about, and of course, how is the happiness and the frustration of the individual while learning with the these technologies.

So, now having understood that ok these are some of the most commonly used affective states and exhibited by the learners, then how would you like to use it? So, the way we can use it that whenever we are developing an affective learning technology. So, what we can do?

We can try to focus more on the monitoring of these states and accordingly we can select the type of the sensors, we can select the type of the ground truth, we can select the type of the machine learning deep learning algorithms in order to recognize and also to adapt to these different states.

So, these are the most relevant affective states that we can use during the affective in the interaction with the affective learning technologies.

(Refer Slide Time: 05:40)



Now, let us try to look at some affective learning technologies and how they have used the affective computing in general. So, first, we will talk about the affective AutoTutor. So, affective AutoTutor was proposed in December, this came in December 2012. But this was a work of previous earlier several years.

And it is said to be one of the first reactive conversational intelligent tutoring system. So, it was a intelligent tutoring system, but it was first reactive conversational intelligent tutoring system, in the sense that it was able to react in an conversational way to the emotions of the individuals or the students.

So, you can see here, I hope that this particular graph or image is visible to you. So, you can see that there is some screen interface that is there in front of a user which could be an student, which could be an individual. And then, this particular machine or the interface that is there, it is tracking 3 different modalities of a particular user.

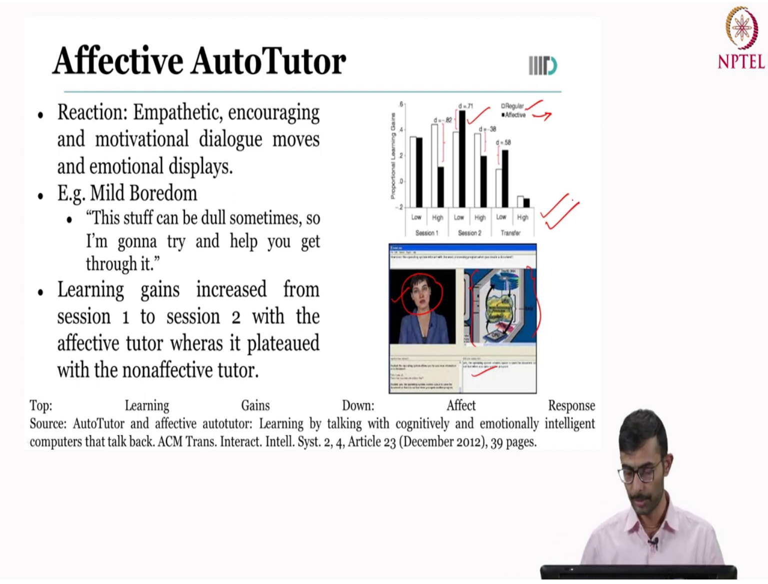
So, it is making use of some camera. With the help of the camera, it is tracking the facial features. It is also looking at the contextual, it is also looking at the contextual cues and then it is also looking at the body languages of the individual. So, this is these are the different cues and the modalities that the different the system is targeting or the system is tracking.

And by tracking all these 3, what it is trying to do? It is trying to detect the boredom, confusion, frustration or the neutral state of the individual because on the basis of these states, the tutoring system would like to do some adaptation. So, apart from the monitoring part, now there is a part of that how to take a particular decision.

So, this is where a decision level fusion algorithm was employed by this tutoring system, in which it clubbed the data from the different modalities, such as the camera or the body languages. And then it applied some fusion system. And after applying that fusion system, it was giving or getting a particular diagnosis of the type of the emotional state that the individual is experiencing.

So, hence that is how we are making use of the different modalities to arrive to a common decision of what is the type of the state among, let us say these 4 that the individual is experiencing.

(Refer Slide Time: 08:25)



Now, once we understood that what state the individual is, the next we would like the system to do some adaptation to it. And the type of the adaptation that happened here in the case of the AutoTutor was that it provided empathetic and encouraging and motivational dialogues and emotional displays in response to the some cues that it tracked or monitored.

So, for example, when it understood or tracked that there is a mild boredom among the individuals or the individual or the user who is interacting with the tutoring system is experiencing some boredom is getting bored. Then, it may give certain type of cues with certain emotional displays as you can see in this image. I hope this image is clear.

So, this particular animated character is providing certain emotional displays and then there are certain emotional dialogues that are being generated here. So, for imagine, I am not sure whether the image is very clear or not. So, this was a particular topic that was being taught let us say about the CPU, RAM, and the central processing unit of a machine.

And then, the agent or the system tracked or monitored that the user is feeling bored about it. So, once the system was able to track that the user is feeling bored about it, then it presented a certain motivational dialog saying, that ok, "This stuff can be dull sometimes, so I am going to try and help you get through it."

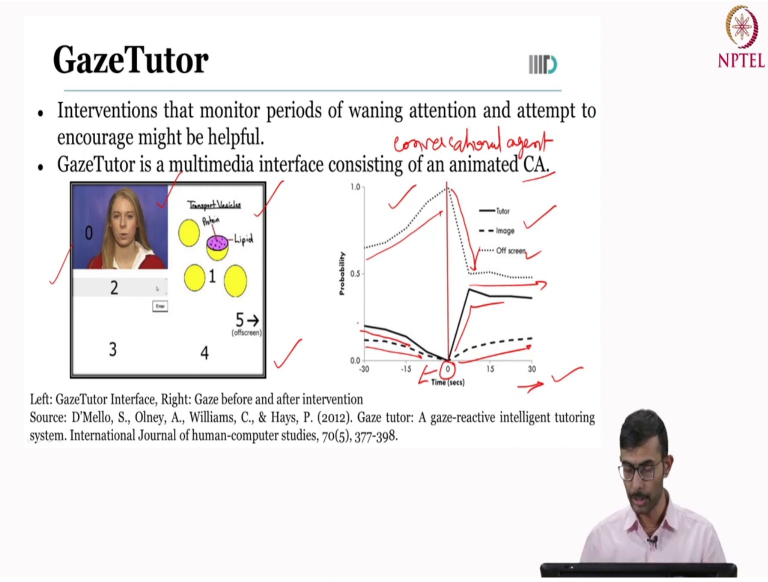
So, this is a simple statement, simple motivational dialog, of course, combined with certain emotional cues as you can see in the animated character, that the system presented which in order to help and motivate the user to improve the learning.

Now, the next question whether the learning improved and as you can see in this diagram that is presented on the top. So, here we are looking at the learning gains from different sessions between the regular and the affective system. So, regular is the one where there was no feedback provided so, regular is the one where there was no feedback provided and affective was the one where the non-monitoring was being done, but also certain type of feedback was provided.

So, as you can see in the diagram that is being presented on the right on the top right corner, that the learning gains. It increased from session 1 to session 2. And so, you can see that these are the learning gains. So, learning gains in the session 1 they were not very high, but in the session 2 they were higher for example.

So, overall, the learning gains we can conclude that the affective tutor was able to improve the learning or improve the learning of the user with the help of this type of adaptive feedbacks. So, that was one very simple example of how this affective auto tutor was able to track the emotional state of a user, of a student and how it was able to adapt and through that adaptation how in gain in learning was achieved.

(Refer Slide Time: 11:40)



Next, we will see another example which is of the GazeTutor. So, in case of the GazeTutor, the basic idea was that if we can monitor the periods of the waning attention and if we can attempt to encourage those at those particular type in those particular periods, then these type of interventions can be very helpful.

And building on this hypothesis the GazeTutor is basically nothing, but a multimedia interface which is consisting of an animated conversational agent. CA is basically the conversational; CA is nothing but the conversational agent. So, CA is the conversational agent.

So, as you can see on the left hand side this is the GazeTutor interface where there is some image about which is being used to present or to display or to teach some particular topic, and then there is some animated conversational agent that is you can see on the left hand side.

Now, on the right hand side what you can what you can see is of course, the idea was that ok, it was able to track your attention period, and then it was able to adapt to that particular attention period or it was able to intervene at that particular point of time. On the right hand side what you can see, that this particular line it represents the attention before or the gaze of the individuals that was before and after the intervention.

So, for example, this is representing that where the this particular area the left side area is representing where the user was looking at when before the intervention and what happened when the intervention was provided. So, if you look at the left hand side then what it is telling you roughly, for example, if you look at this thing the x axis of course, is telling you the time and the y axis is telling you the probability where the individual is looking at.

So, for example, if you look at just one the off screen time here so, off screen time you can see, the off screen time is gradually increasing with as the time progresses. So, what it means that individual is neither looking at the screen, in the individual is not looking at the tutor, individual is not looking at an image, right.

So, basically, this is the tutor, this is the screen interface and this is the particular image that is being presented, but the individual user is not the user is not looking at either of them and the off screen time is gradually increasing. The individual is not paying attention at all. And then a certain intervention was provided.

Now, if you look at this as certain intervention was provided, then what happened? Then, gradually very quickly the off screen time it started to reduce, and then gradually, it came at a particular level it got sustained at a particular level where the individual's probability of looking outside the screen was around 50 percent 0.5.

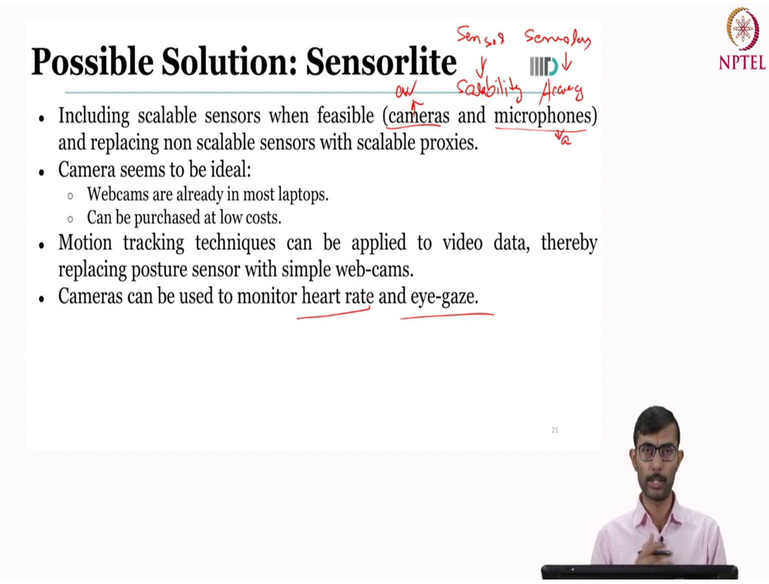
Similarly, the same type of analysis can be looked at the tutor and the image. So, you can see that this is the decline that you can see with respect to the tutor and the image. So, before the intervention the individual was not looking at the tutor or the amount of time that it was the, as the time progresses the individual started not looking at the tutor. So, this is the tutor, not the agent. And then, this is the image.

So, basically, the individual user was neither looking at the tutor not nor looking at the image and this is represented by the fact that the time, the probability with which it was looking at there started getting reduced. But then, as the intervention was provided, after the particular type of intervention you can see that the probability with which the user was looking at the tutor or even for that matter to the image, it started increasing.

Of course, it could have increased a bit more, but nevertheless it was much better than without the intervention. So, that is how the GazeTutor was taking care of the waning attention period, and was able to provide some intervention at this particular point of time.

(Refer Slide Time: 15:39)

(Refer Slide Time: 15:48)



So, now we will look at the affect-aware games. So, we saw so far that if we go for a full sensor based approach, dedicated sensor based approach, then there is a problem of the scalability, right. So, let me just write it down for you.

So, if you are going ahead sorry; if you are going ahead with the sensor, then dedicated sensors they will require, you will face a problem of a scalability. And if you are going sensor less, then you will have a problem of accuracy. So, maybe the accuracy you are somehow making a trade off with the accuracy, sensor less, right.

So, now, so there could be a possible solution where we are we where we which we are calling it as a sensor light solution. So, sensor light solution means we are going to use the scalable sensors whenever it is feasible, such as for example, you know cameras and the microphones they are very very widely available sensors which can allow you to capture them audio visual modalities and only the audio modalities. And that is one approach.

And then, at the same time what we can do? The non-scalable sensors can be replaced with the scalable proxies. So, for example, if you are trying to make use of camera can be very good option here; and for example, because once that it is already available in most of the laptops. Similarly, for example, it can be purchased at very very low costs.

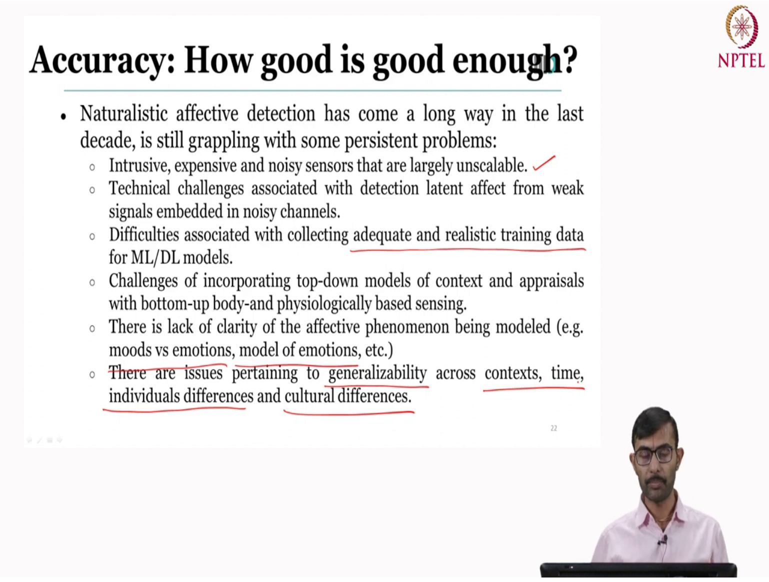
And previous research has already shown we have talked about it that you can use the camera as well to for example, record and monitor the heart rate, and the heart rate variability and related features. So, this seems like an excellent solution. And other thing for example, that you can do, you can also simply look at the webcam data and you can apply the motion tracking techniques on it.

So, for example, you can do the posture analysis, gesture analysis all with the help of the 2D visual data. Of course, you will need to work a bit more on the software side, but working on the software side is a bit more easier, it is more accessible than you know working on the hardware and trying to make the hardware scale.

So, so that for in this way, you know like there is one type of possible solutions. Other, we already talked about that you know like the camera can only, not only be used to monitor the heart rate and the heart rate variability, it can also be used to monitor the gaze patterns of the eyes. And hence, it can also replace to certain extent the eye tracking devices and the sensors that we have.

So, that is the conclusion of this thing. If you have the possibility of going making use of the dedicated sensors, go ahead. Many times, your application domain will not allow it. Then, you may want to use the existing sensors which are there available in the system to track the behavioral data or for example, you may want to replace the dedicated sensors with the sensor proxies and hopefully it may work wonders for you, ok. Sorry. So, that was about the sensor and what to do with the sensors that are available to us.

(Refer Slide Time: 19:08)



Now, let us talk about the accuracy as well. So, in general, there is a we know that you know the naturalistic affective detection has seen a lot of research and it has been improved a lot. But of course, there are lots of problems that are associated still with the affective computing domain and in general the research and the development that is there.

One thing we already talked about it there, most of the time the sensors that we require they are intrusive. Many times, it could be expensive as well, many times it could be noisy as well and more importantly they are not scalable. Other thing that we see that you know technical challenge, if you look at talk about the technical challenges, then the detection itself can suffer from the weak signals that are embedded in the noisy channels.

So, many times for example, you are trying to look at a particular data and then, but that data itself is surrounded by lots of noise around it. For example, maybe you are trying to capture the emotions in the voice, as simple as that. You are trying to capture the emotions that are there in the audio modalities of a user, but now we are no more in a lab setting, we are in a naturalistic setting where there is a lot of noise around it.

And hence, your target user's voice data is getting embodied is getting surrounded by lots of noises that are around there. And hence, it really becomes very very challenging to segregate that data of the user, voice of the user and then and do some analysis on the top of it.

One thing also that we have seen even though we did not talk about a lot in detail about the machine learning and the deep learning algorithms itself. But most of the time what we want to do whenever we are talking about the fatigue computing, it is going to rely heavily on the machine learning and the deep learning algorithms.

And whenever we talk about the machine learning and the deep learning algorithms, it is no surprise that they need a lot of adequate and realistic training data in order to make lot of sense. And many times, what happens that emotions data associated with the emotions are maybe very very difficult to even capture to annotate. And hence, in turn our machine learning and the deep learning models may suffer, and their accuracies may suffer.

So, other thing for example, that may happen that whenever we are talking about the affective computing, most of the time we are just looking at the transient emotions. But now if you want to incorporate the context and the appraisals as well into it, then this becomes a bit troublesome because you know if you want to incorporate the context and the appraisals.

And let us say you know the users beliefs desires and intentions around it, then it can be become very tricky because then you will have to be able to you know track lots of different things, and which may be very very difficult here. In general, you know whenever we are talking about the affective computing.

Also, there is a common mistake that the researchers do that the developers in the community do and they are doing, that many times they are taking a replace using one for the another and the another for the one. So, for example, whenever they are talking about the model of the emotions, they are not able to discriminate properly between the categorical models or the continuous models.

Similarly, for example, many times the mood versus emotions is not being differentiated. So, you may refer to mood which is you know long term emotions or maybe you are referring as a transient emotions, but you are just you know using the one for the another. And hence, you are notable to differentiate these things, and hence there is a lack of clarity on your side.

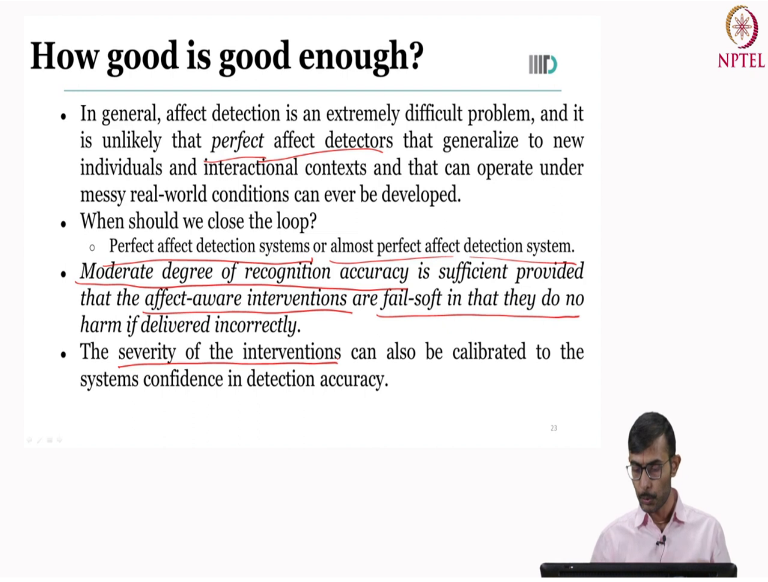
Nevertheless, even after this if you are able to do some recognition and you are able to make use of some monitoring, then or what happens? That the generalizability itself it becomes a big, big issue.

And generalizability in general for the machine learning and deep learning algorithms itself is a problem, but more so, with respect to the emotions and then the affective computing it is a bigger problem. Because you may want to look into the individual variability, we talked about this.

You may want to look at the cultural variability, cultural differences. For example, the death of an individual, at the time of the death of an individual the way it is being expressed in Indian communities is very very different from the way it is being expressed you know the lamentation or the sorrow that is being expressed in the outside community.

So, of course, there are lots of cultural differences that are surrounding it. Of course, then context, time, etcetera, they also play a lot of role. And so, generalizability becomes a big, big issue on especially we will talk about the affective computing models.

(Refer Slide Time: 24:15)



So, the question that in general we want to ask now is this, that ok, what type of accuracy or to what extent an accuracy is a good accuracy with which you know we can go ahead and deploy the system and we can make it work? So, in general, we already saw that affect detection itself is a very tricky problem. And we can very confidently say that ok, it is going to be very very unlikely that very soon we are going to have a perfect affect detectors.

You know that are not only able to you know do a 100 percent classification of the emotions of a user in real time. But also, they are able to generalize well to the new individuals to their international context and in you know very noisy situations that are around us.

So, it is not going to happen anytime very soon. It will require lots of advancements, not only on the hardware side, but also on the software and the machine learning and the deep learning side from us. So, the idea is ok, when should we start taking the information that we are getting from the emotion recognition systems and build on the systems that can adapt to it or in general when we want to close the loop.

So, there could be two possibilities now, whether we can go wait until there is a perfect affect detection system in order to build the adaptive system on top of it, which is going to work perfectly fine. Or we can just go for a not so perfect affected affect detection system and we can try to make it work.

So, the thing is here. Since, it is a very very tricky problem, it may take a lot of time and the resources to arrive to that situation. So, even if we are able to get a moderate degree of recognition accuracy. And what is a moderate degree of recognition accuracy? For the emotions, that depends on them situation to situation and domain to domain and as peruse cases. We will talk a bit more about it.

But the moment we are able to get or you are able to get a moderate degree of recognition accuracy, we believe that should be sufficient to, so to create the affect-aware interventions. So, affect-aware intervention means adaptive interventions. We just saw some of the examples, some of them.

But of course, we have to take into account the fact that there should be fail-soft. Fail-soft in the sense that they should not do any harm. Let us say even if they are doing some adaptation, which is based on the incorrect classification. So, that is a very tricky suggestion that you want to look into it.

And of course, now for example, when we say that the moderate degree of recognition and then accordingly the severity of the interventions. So, of course, you want to make an adaptive system, but since you are making an adaptive system which is not entirely, which is based on not so perfect affect detection systems. So, you may not want to put a lot of hard confidence in the severity of the interventions.

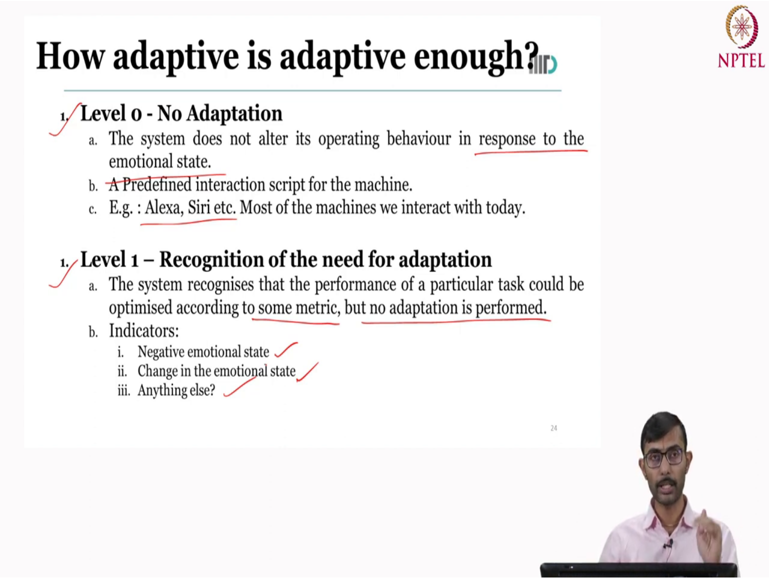
So, you may want to take it up in the pinch of salt, that ok, I got a particular classification, I got a particular emotional state of a user based on whatever sensors or the models or the software that I have deployed. And there is a possibility that it may not be so correct. Hence, the adaptation that I am going to do is going to be accordingly very not so severe and could be a moderate.

And then, it can be calibrated to the extent, you know that I am able to put the confidence in my detection accuracy. So, for example, imagine I am I can come in my particular use case, I am able to use make use of the dedicated sensors, dedicated hardware’s, and I am able to use instead of the art algorithms, and then I know the detection accuracy is good or is almost perfect in my case.

So, so then maybe you can put more confidence in the system and accordingly the adaptations that you are making, maybe you can put more confidence in the adaptations as well, and then accordingly so on so forth, right. So, this is a very good important thing to understand that you need to take into account the system that you are building.

And then, the use case that you are building, and accordingly, you need to define that ok, what could be a moderate degree of recognition for you, what could be the severity of the interventions that you want to go ahead. But bottom line is you may never want to wait for a perfect affect detection system in order to start building adaptive system, ok. So, that I hope is a bit clear.

(Refer Slide Time: 28:56)



Now, having talked about ok, the sensors, having talked about that what is a good accuracy, what enough what accuracy is good enough. Now, let us talk about a bit on the adaptiveness, right. So, the question that we want to answer ok, what should be the adaptiveness of the system?

And this is a very tricky question again. And the severity of the adaptive interventions as we said already is determined on the basis of the confidence that you are able to put into the system. But, let us see what are the different levels of the adaptation that we can have in the system.

To begin with, for no surprise, we can have a level 0 adaptation which we also call it as a no adaptation at all. So, in the no adaptation at all, what it means that we do not expect the system to alter its behavior in response to the emotional state as simple as that you know.

Whatever is going to be the emotional state, we are going to monitor it, but we are not going to do anything about it. We are simply going to take that information you know for some other analysis purposes, but we are not going to let the systems behavior altered by the response and response to the emotional state. And that is what is happening for most of the systems that we have.

And for this what happens you know, a pre-defined interaction scripts mostly you know is used for the machines, for the services that we used, we just talked about the gaming examples. So, you know when you are sad the non-playing characters also around they show a sadness, but all this is very very pre- defined script. And it does not really know; first thing it does not really know that what your emotion is, and even if you knows it does not, it decides not to do anything about it.

So, this type of system is where there is no adaptation as all happening and most of the machines that we are happening interacting with the today or the services including for example, you know the voice agents such as Alexa, Siri, most of the machines are like that. Their adaptation is not based on our emotional state and many times our emotional state is also not being monitored. So, that is the level 1, sorry level 0.

Then, comes the level 1. So, the level 1 base is basically you know what it does, it tries to monitor your emotional state number 1, and then, it also tries to recognize that ok there is a need for the adaptation at a particular time. So, it tries to identified the time of the intervention or the need for the intervention. And, but of course, it just does that. It simply tries to identify, identify it is that ok there is a need for the adaptation, but it does not perform any adaptations at all.

And for example, you know the of course, all the when how will we identify that ok; now there is a need for the adaptation, it could be on the basis of many different metrics and some metrics of indicators could be like this. For example, the system is able to, system or the service that you are interacting with is able to understand that ok, you are experiencing a negative emotional state because of your voice or through any different morality.

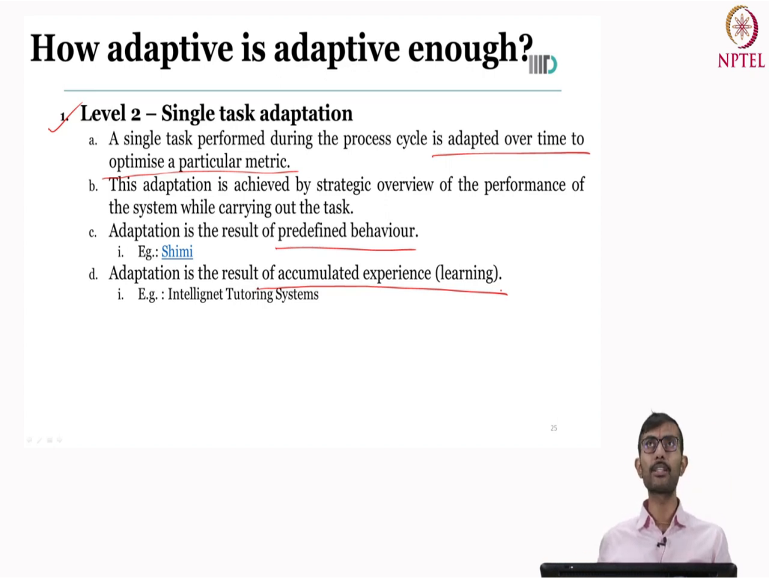
And now it knows that ok. Since, you are experiencing a negative emotional state you are feeling low, you are feeling sad, you are feeling bored and so on so forth. So, what it means that ok there is a need for the adaptation here. Similarly, whenever there is a change in the emotional state, maybe you can just look at this metric.

The systems can look at this metric, that whenever there is a change in the emotional state of a user, maybe the system also needs to adapt to that and that is the metric that you are; of course, you are not doing the adaptation as we already talked about it. That, no adaptation is being performed at this stage, but you are able to identify the need for the adaptation.

Negative emotional state, we already talked about it. Changes in the emotional state, whenever you change your emotional, let us say from one to another maybe you are feeling happy and suddenly you started feeling sad. So, then maybe the system will get an alert oh, the individual was feeling happy and now the individual is feeling sad. There is a change in the emotional state and maybe this is a right movement to do some adaptation. But in the level 1, of course, no adaptation is going to do.

And then, you can think about like lots of other different metrics. I will let you explore and think that what could be the other metrics on the basis of which you may want to do some interventions or adaptations, right. And so that is the level 1 system. So, level 0, no adaptation. Level 1, it just recognizes the need for the adaptation, but in general there is no adaptation. As of now, yet we are not doing any adaptation, ok.

(Refer Slide Time: 33:38)



So, now we go to the level 2 which is a bit more fascinating, which is the single task adaptation. So, single task adaptation means what? What it does? A single task that is being performed during the entire cycle is adapted over time to optimise the particular metric.

So, we already have some metric using which we are saying that ok; we have first we are tracking the emotional state of the user, we also have a metric or indicator which is telling us that we are recognizing the need for the adaptation. And then, we are also in the level 2, now on the top of it we are also doing the adaptation in a single task.

And this adaptation can be you know we may want to look into that ok, what is the metric that we really want to improve in terms of the performance, performance metric and how can be it improved by doing a particular type of adaptation, just to make example very clear.

For example, maybe you know you are looking at let us say, let us say you your aim that ok you are building a conversational AI agent and then your aim is to make the user happy. So, you know at the end of this interaction, the user should be feeling happy, user should remain happy, user should feel happy about the interaction of once the interaction is over.

So, whenever the user you are monitoring the emotional state of the user, and whenever you see that ok, whenever you experience that ok user is feeling a bit low, then what you want to do you may want to do certain type of adaptation which will make the user happy. Because you your aim was to make the user happy and so you will do a particular type of adaptation that is going to make the user happy.

Similarly, for example, if you talk about the affective learning systems. So, where your aim is to improve the learning, then you may maybe what you want to do, you are tracking the emotional state of the user and your continuous aim is to keep the user engaged or keep the boredom of the user as low as possible and then so on so forth.

So, then accordingly, the type of the adaptation that you would like to do will be to address that particular goal. Maybe you want to make the user happy, maybe you want to make the user engaged, maybe you want to make the users boredom low and then so on so forth. So, I hope that this is what, this is what it is clear. That, you are able to adapt with respect to the performance metric that you have kept as a goal for yourself.

And many times, this adaptation you know it itself what can what can be the adaptations could be there? It could be the result of some predefined behaviour. So, you are saying that ok my user is feeling sad, I am going to make the user happy. But what will I do to make my user happy, now this is a question. And this particular type of behavior, in this particular system, in the level 2 system is the result of a predefined behaviour.

So, for example, we already saw, we already looked at the example of the shimmy robot. So, in the case of the shimmy robot for example, what happened? That the shimmy robot was able to you know play some music as per the emotional state of the user, but these particular type of music or the behaviour of the gesture of the robot was predefined by the script that the developers have already built in.

Similarly, adaptation can also be the result of the accumulated experience or the learning that happens over the period of time. So, for example, while it is not happening to the extent that we want it to be done, but in the intelligent tutoring systems what can happen that you may you know that the entire experience of the user’s interaction with the system over a period of time. And then, you can learn from that particular experience and you are able to adapt on the basis of that.

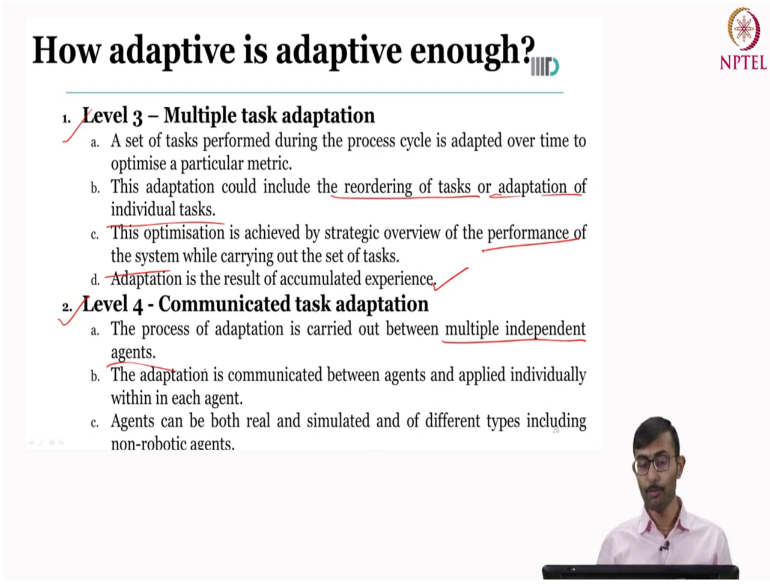
And then you know for example, what type of adaptation will work for this particular guy. So, very simple example is the you know when for example, as a teacher, when the teacher is taking a class in the of for certain students, then teacher knows by the time. You know in the beginning maybe the teacher may not have a very good idea that you know what, how should I teach a particular topic to a particular student in order to make him or her understand.

But over the period of time, you know the user know the teacher knows, a good teacher at least knows that ok, I need to address this particular problem of a, this particular student in such a way, so that you know like it helps that particular student in a specific way.

So, the type of the adaptation that the teacher does for a specific student is different that it does for the other student. And this is the result of the learning that teacher has done over the period of time. And this is what we are envisioning here. That, if the system can learn from the accumulated experience of interaction with the users.

Then it can improve the, it can create the, it can does the adaptation not on the basis of some predefined behaviour, but on the basis of the learning that it is doing. And it is going to be very very adaptive and personalized.

(Refer Slide Time: 38:48)



So, that is the level 2 adaptation which is a single task adaptation and of course, keeping a particular performance metric in the mind. Now, level 3 is a next level adaptation. In the level 3 what happens, that rather than targeting a rather than doing the adaptation in a single task, where for example, you know maybe in the case of the shimmy robot, maybe the robot was only able to do the adaptation in its voice or maybe was only able to do the adaptation in its gesture.

Now, we are talking about that set of different tasks that are happening during the process cycle, can be adapted over the time in response for to optimize a particular performance metric. And this adaptation can be of many different types. So, for example, adaptations could be that there are multiple tasks that the agents are doing, that your machines are doing, that your services are doing, you can simply do the reordering of the tasks. Or you can simply do the adaptation of the individual tasks that are happening, but in parallel.

So, for example, a good example of this thing would be that if you want to say that ok, you want to do the reordering of the tasks based on the adaptation. Maybe you know for example, you are trying to create an intelligent tutoring system here, and the tutoring system you know the movement for example, the user logged in to learn a particular concept, maybe the system felt ok, the individual looks a bit you know like low on the energy today.

So, if the individual looks a bit low on the energy, I am going to you know teach maybe topics that are easy to follow first, rather than you know start with the hard topics. And maybe I am going to teach in a way that is you know at a very very basic level. For example, multiple ways that this type of adaptations are being done.

And so, your reordering also and the multiple tasks are being you know taken into account. So, you are looking at you are not only adapting your teaching style, you are also changing the content that you want to teach for example.

Or in the other way you are simply adapting multiple tasks, but you are not doing any reordering of that. For example, imagine that you are interacting with a chat bot, a conversational bot, and the conversational agent is not doing the reordering of anything, it is following the order that it is supposed to follow, but maybe you know it is not only adapting its gestures in response to your emotional state. But also, it is adapting its voice as well in response to your emotional state.

So, there are two tasks at least you know. And maybe on the top of it maybe you know the task that it is supposed to perform, actually, maybe you have a problem with the bank or something like that, it is also able to do it in a way that really praises you. So, there are lots of different tasks that are happening, lots of different processes that are happening, and the adaptation of all these processes or multiple processes are happening in at the same time in this level 3 type of adaptation.

And of course, nevertheless, no matter whether you are adapting a single task, multiple tasks, whether you are reordering the tasks, whatever the type of the adaptation, this all the adaptation has to keep in mind the performance of the system. So, basically, you have a particular goal in mind, you want to make the user happy.

You want to make the user feel fulfilled, you want to make the improve the learning of the user whatever. You have already all these goal predefined. And all these adaptations are happening as per the goal, broader goal that you have set for your system, right.

And in this case, of course, you know, rather than having a predefined script, what you simply have, you simply have adaptation that is the result of the accumulated experience. So, it is very very personalized and very very customized for each different user. So, the adaptations are going to be different for each different user depending upon their likes and dislikes. So, that is quite interesting.

Now, in the level 4 adaptation, it is very much like the level 3, but there is a critical difference that the process of the adaptation is carried out between multiple independent agents. So, what happens? That till level 3, we are assuming that we have a system or a service or a machine where there is only one agent with which we are interacting with. And that agent is adapting to our task or is not adapting or is adapting multiple tasks that the agent itself is performing.

Here, we are saying that we are interacting in a multi-agent setting or we are interacting with the multiple services at the same time. And all these multiple services, multiple agents, they are talking to each other and they are saying that ok you know, like you adopt this, you adopt this, and I will adopt this, and let us collectively make the user feel good about the entire thing.

So, for example, you know, maybe you are playing a game and the game there are multiple characters. So, when there are multiple characters, rather than one character adapting to you, all the different characters, they are adapting to you, but they are doing in sync by communicating with each other. So, this is really fascinating because now, the your user experience is being looked at holistically and comprehensively, and maybe it can provide a better experience overall.

And while doing so, of course, what the agents can be do, they can communicate the different adaptations and they can be applied individually within each agent. So, as I said agent A can say to agent B, service A can say to service B, that you know, we have to do this, this, this and service A will do this type of adaptation, service B will do this type of adaptation.

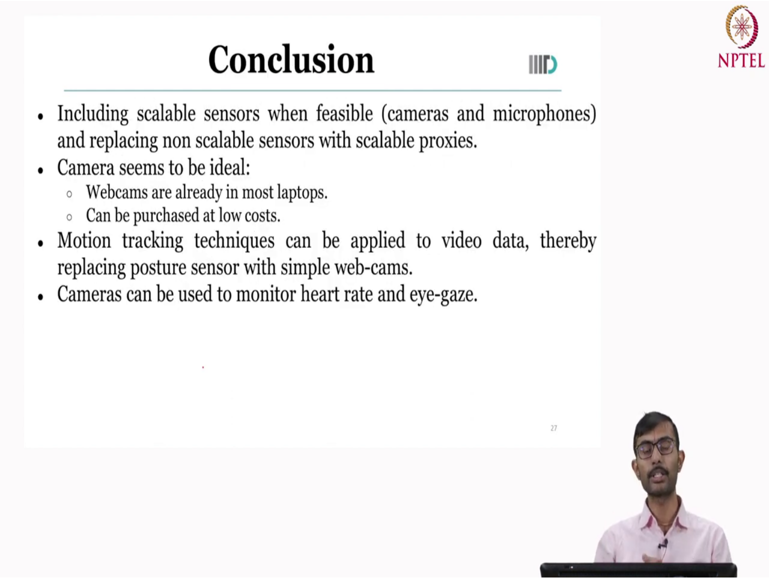
Similarly, you know, agent A, agent B, machine A, machine B or whatever different types of things. So, for example, you know, there are different, let us say, in the gaming itself, I will there are two different characters that are around you, they are supposed to be helping you in, I do not know, finding some treasure and then.

So, agent A is going to, you know, maybe say that ok, maybe the user is not able to find it, let us help the user find it. So, agent will say, ok, maybe I am going to, you know, clear the path for the user, while let us say, you know, you take care of the, I do not know, enemies that are there on the path. So, different things, right. So, different, of course, depending upon the capabilities of the agents or the services.

And as I said again, the creativity is the only limit here again for you, on the type of the adaptations that you can do for your machine, ok. So, again, when we are talking about this multiple agents, this multiple agents can be both real and simulated and of different types as well. So, you, we are talking about the agents in the games, we can also have one agent in the embodied agent, in the animated agent, one robotic agent and, and then so on so forth.

So, basically, all the different types of agents or the services that we can envision, they can work together, imagine you have a robot at your home, you have a Alexa also, you have a Siri also, and they all are talking to each other in order to make you feel happy, I mean, that would be really nice, for example. So, that is what is known as the communicated task adaptation.

(Refer Slide Time: 45:45)



Now, in conclusion, when we looked at the, so this is the, let us look at the conclusion now. So, when we looked at the open issues here, we already saw that we can use the scalable sensors, whenever there is a feasibility of it, such as in the case of the cameras and the microphones. And we can also replace the non-scalable sensors with the scalable proxies, which can look at the latent behaviour, such as behavioural data. The one, that we, for example, the way we capture it from the keyboard, typing and so on so forth.

Camera is a very very ideal choice, because it is already available in all the laptops systems, and if not, then it can be also purchased at very very low costs. Motion tracking techniques can be applied to the video data. Hence, you know, like it can also enable this kind of tracking, and so this is, these are some of the things that you want to look into the, when you are looking at the conclusion of it.