NPTEL

NPTEL ONLINE CERTIFICATION COURSE

Introduction to Machine Learning

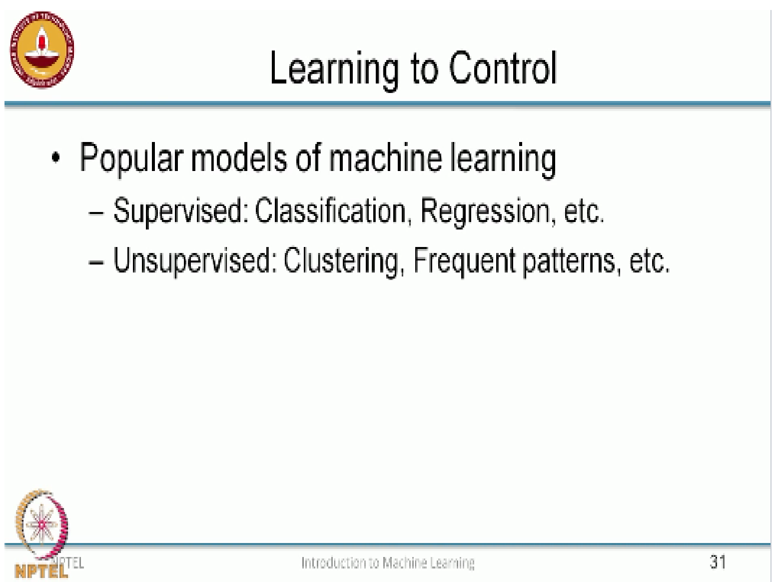
Lecture 4

Prof: Balaraman Ravindran
Computer Science and Engineering
Indian Institute of Technology Madras

Reinforcement Learning

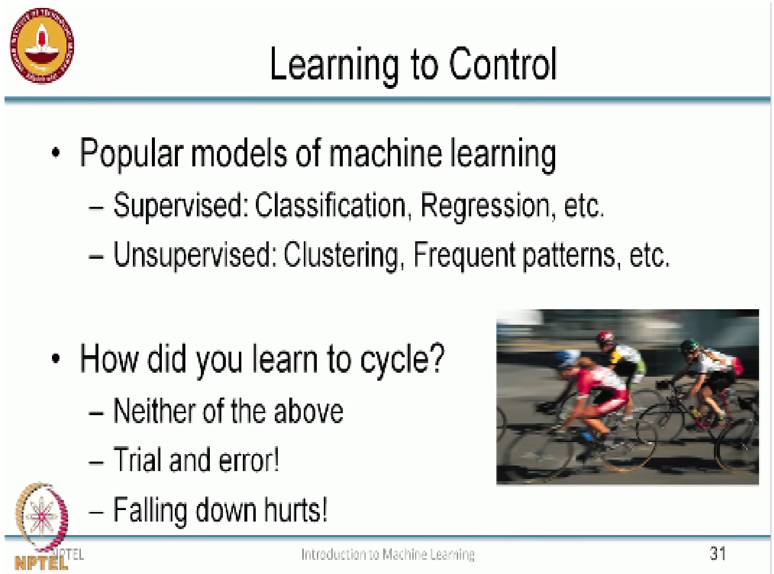
Hi and welcome to this module that reduces reinforcement learning so far we have been looking at popular models of machine learning such as supervised and unsupervised learning the supervised learning we looked at the classification in the regression problem and in unsupervised learning we looked at clustering and frequent pattern.

(Refer Slide Time: 00:32)



So on so forth and I have a question for you so. So how did you learn to cycle was it supervised learning or was it unsupervised learning right there is really no one telling you how you should cycle right.

(Refer Slide Time: 00:45)

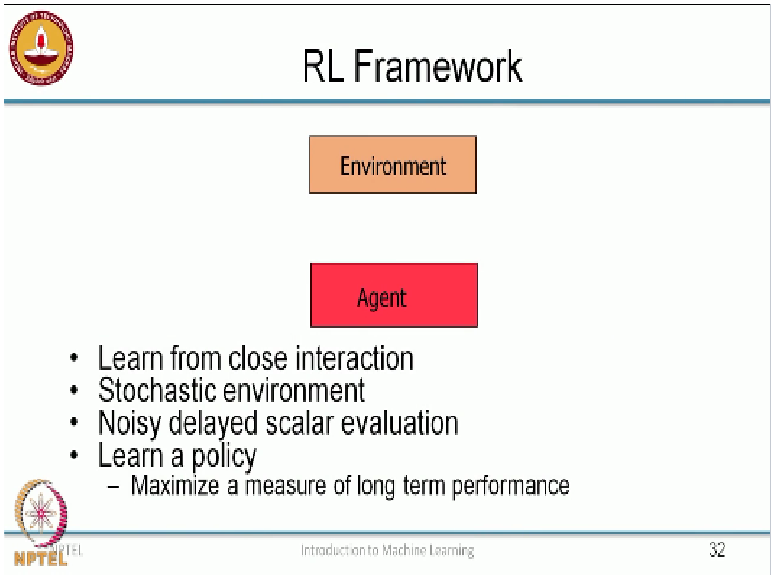


I mean how much how many pounds of pressure you should put with your left foot and what angle should be leaning and so on so forth and if you think of it as a supervised learning problem that is how it should be and it was not completely unsupervised because it is not like you just watch people cycling and then figure out what the pattern that you should move in order to cycle and then you just magically got one cycle of chartered cycling right.

So what is the crucial thing here right there is a trial and error component right so you have to get on the cycle right so you had to try things out yourself before you could learn how to cycle in a acceptable manner right so you have some kind of feedback it is not completely unsupervised right there will be somebody cinder is standing there and if you learn to cycle as a kid there was somebody standing there and clapping and saying hey great good job go and go on go on or something like that and of course falling down hurts.

So you know that right so there is some amount of trial and error component and that is feedback that you are getting from the environment so this kind of learning where you are learning to control the system through the trial and error and the minimal feedback is essentially what reinforcement learning is it is a mathematical formalization that captures this kind of learning is what wherever to as reinforcement learning right. So in the RL framework you typically think of a fun learning agent.

(Refer Slide Time: 02:12)



That we already looked at learning agents X it could be the supervisor learner or it could be none supervised agent in this case you have a reinforcement learning agent that learns from close interaction with an environment that is outside the control of the agent right the oral agent learns from close interaction with an environment so what do I mean by close infraction here is that the agent senses the state in which the environment is right and it takes an action which it then applies to the environment.

Which causes the state of the environment to change so thereby completing the interaction cycle so the agent senses what is the state of the environment so if it is a cycle it is going to sense what angle is the cycle tilting in at what speed I am moving forward right and on what speed I am falling and so on so forth all this constitute the state of this system state of the environment the agent is going to take an appropriate action which would be okay.

Lean to the right or push down with your right leg right and then this action is then applied to the environment and that in turn changes the state of the environment right so the agent learns from such close interaction with the environment and we typically assume that the environment is stochastic so every time you take an action you are not going to get the same response from the environment so things could be slightly different right so there might be a small stone.

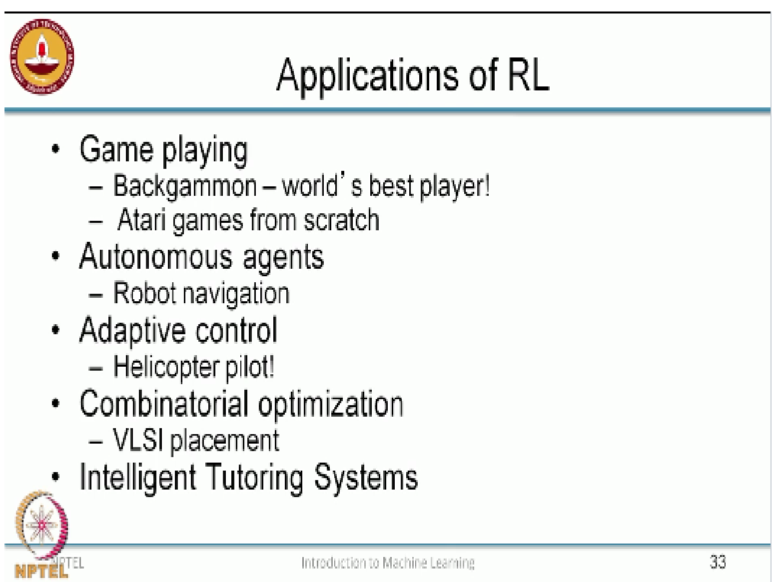
in the road that you did not have the last time you went over this place and that for what was a smooth ride could suddenly turn bumpy and thus once of course I mean so you know that cycling always has some amount of noise and then you have to react to the noise so apart from this interaction the mathematical extraction also assumes that there is some kind of an evaluation signal.

That is available from the environment that gives you some measure of how well you are performing in this particular task right if you remember we needed to have an evaluation measure for every task and we are assuming that this comes in the form of some kind of a scalar evaluation from the environment it could be somebody clapping and saying that here you are doing well or it could be falling down and getting hurt so all of this would be translated to some kind of a numeric scale.

And that's the mathematical abstraction that we make right so the goal of the agent is to learn a policy which is a kind of mapping from the states that you sense to the actions that you apply so as to maximize a measure of long-term performance so I'm not just interested in staying upright for the next two seconds but I am really interested ingesting from point A to point B so I have to make sure that I stay balanced throughout the entire duration of the right so this is the basic idea behind the reinforcement learning problem so each reinforcement algorithm the goal into learn a policy that maximizes some measure of long-term performance right.

So there have been many successful applications of reinforcement learning so one of some of the marquee applications come from the domain of game playing like with many classical approaches so backgammon is a board game.

(Refer Slide Time: 05:18)



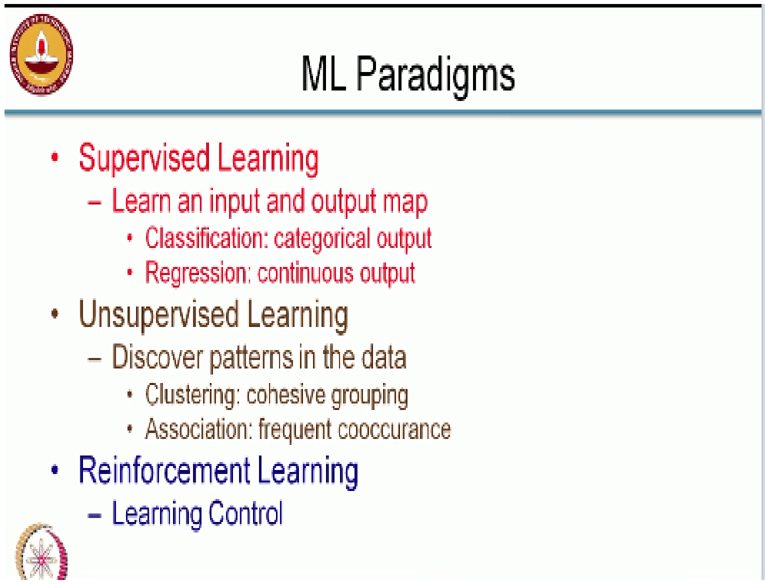
Based on die rolls if people have not familiar that Kevin is similar to the game Pluto but it's also got a rich history people have been playing it for several centuries and there are even world championships in backgammon and the world's best player of backgammon is actually reinforcement learning engine so notice that I did not qualify it saying the world's best computer player or anything so it was the world's best player and that managed to beat the world champion in backgammon over tournament right and so.

More recent vintage so people have also gotten reinforcement learning agent to play at the video games Academy video games from scratch so the input to the system we are like pixels from the screens right and the output from the system where the joystick controls and they managed to play this games from scratch right and so in autonomous agents like in robots and other autonomous agents reinforcement learning is almost always the learning algorithm of choice and so in adaptive control and one of the again very prominent success stories of reinforcement.

Learning is this helicopter pilot rather was initially trained by entering at Berkeley and later at Stanford we were trained reinforcement learning algorithm to fly helicopter and at near human level competence right and there are other applications where people have looked at applying within common atotaloptimization problems following really hard optimization problems and also in personalization and in adaptive systems like intelligent tutoring systems right.

And so to wrap up the set of introductory modules just wanted to recap the different machine learning paradigms that will be bedcovering in the course so the first one we will be looking at is supervised learning where we will be looking at learning in input-output map.

(Refer Slide Time: 07:31)



And so the classes the tasks that we look there or classification where the outputs that we are looking to predict or categorical outputs like yes or no or blue or red or by is by a computer or not by a computer and the second supervised learning problem we look at is a regression where the output is continuous output and the second class of problems we look at are unsupervised learning problems where we are interested in discovering patterns in the data.

Not necessarily in predicting a specific output and the canonical class we look at here or clustering but we are interested in finding cohesive groups in the data and also Association rules where we are interested in finding frequently occurring patterns right and the third paradigm which we will spend very little time on is reinforcement learning where you are interested in learning control or learning to control system based on a animal feedback so from the next module onwards we will start looking at taking a little bit more mathematically rigorous look at machine learning.

IIT Madras Production

Funded by
Department of Higher Education
Ministry of Human Resource Development
Government of India

www.nptel.ac.in