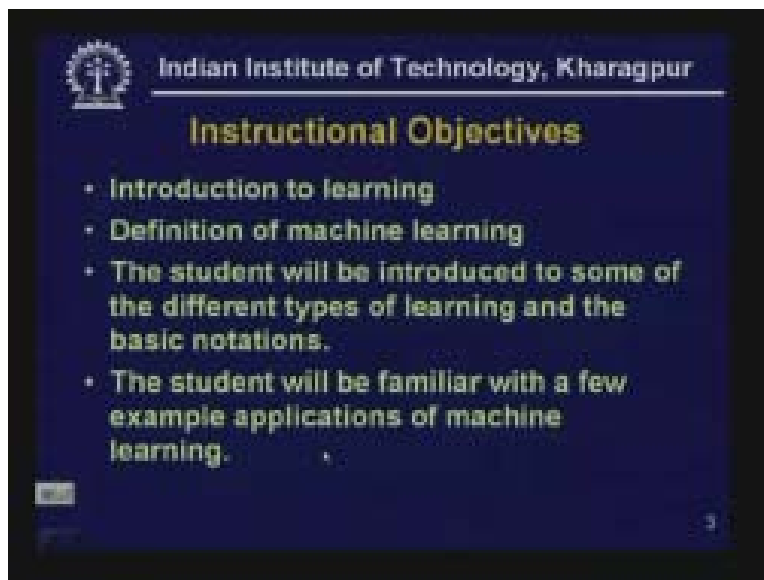


Artificial Intelligence
Prof. Sudeshna Sarkar
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur
Lecture - 32
Introduction to Learning - I

In today lecture we start with the topic machine learning. We will have seven lectures of this series. Today we will give introduction to this topic. The instructional objectives of today's lectures are the following. We will first look at what we mean by machine learning and several definitions of machine learning.

(Refer Slide Time: 01:23)

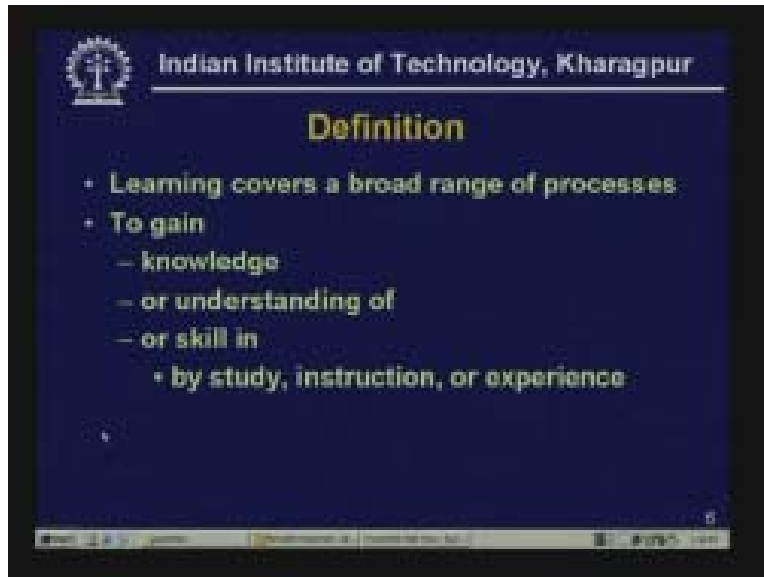


In this class the student will be introduced to different learning frameworks and then we will introduce some of the basic notations. Also the student will be made familiar with certain example applications of machine learning. Specifically we will give an introduction to the type of learning which we call concept learning. In this context the student will look at certain things; what are the concept of features used in a learning problem.

We will look at what we mean by a hypothesis base or a set of hypothesis and what we mean by a hypothesis that we are trying to learn. We will introduce what we mean by a training set and a desk set. We will also talk about the instance space. Also we will briefly introduce the notion of inductive bias. On studying this lesson the student should be able to formulate a given concept learning problem. Given a problem they should be able to identify possible features that may be relevant to the problem and they should be able to get an idea of the hypothesis base that they need to consider.

In subsequent lectures we will look at different algorithms and different types of representation issues. Let us look at the definition of learning. Learning covers a broad range of processes.

(Refer Slide Time: 03:11)



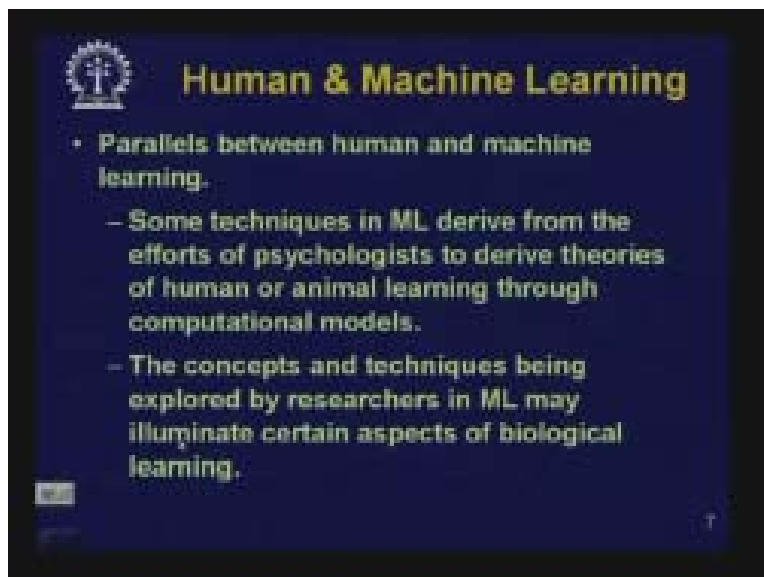
Indian Institute of Technology, Kharagpur

Definition

- Learning covers a broad range of processes.
- To gain
 - knowledge
 - or understanding of
 - or skill in
 - by study, instruction, or experience

Learning means to gain knowledge or an understanding of skill in something. When we say we want to learn then we want to gain knowledge or understanding or expertise in solving some problems. And this expertise or this knowledge can come by looking at the examples or by gaining example experience or this knowledge can come by studying the problem or by being told by instructions.

(Refer Slide Time: 03:59)



Indian Institute of Technology, Kharagpur

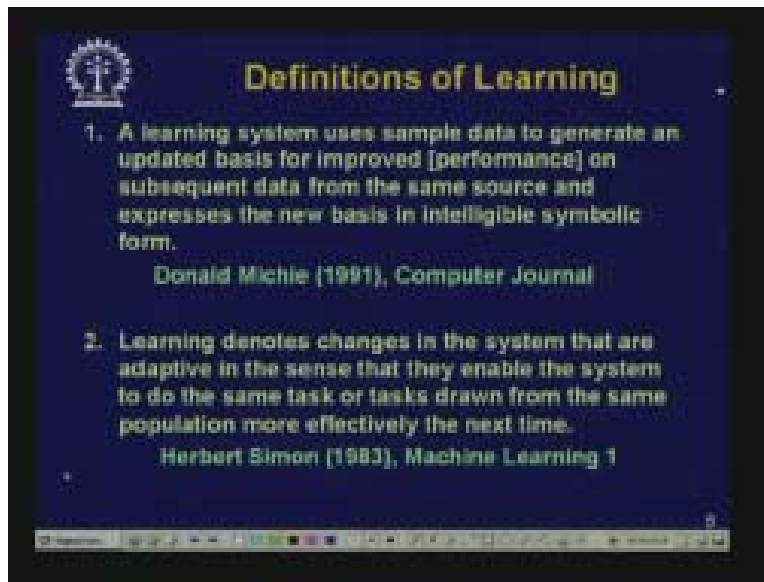
Human & Machine Learning

- Parallels between human and machine learning.
 - Some techniques in ML derive from the efforts of psychologists to derive theories of human or animal learning through computational models.
 - The concepts and techniques being explored by researchers in ML may illuminate certain aspects of biological learning.

There are several parallels between the process of human learning and artificial or machine learning. Some techniques in machine learning derive from the efforts of psychologists to derive theories of human or animal learning through computational models.

The field of cognitive psychology has dealt with in trying to understand how humans learn. And in some cases they have also tried to identify certain computational models to model the way humans think and the way humans learn. Machine learning on the other hand has dealt with certain techniques some of which I inspired by human learning techniques as derived from cognitive psychology, others by different symbolic techniques which are efficient to do by machines. But it is conceivable that the concepts and techniques being explored by researchers in machine learning may actually help understand certain biological processes. So there is a lot of cross partilisation between the wheels of cognitive psychology which tries to understand animal learning and the field of machine learning, their objective is to solely learn by the help of machines. Before we formulate to define learning let us look at few of the definitions that people have put forward on machine learning.

(Refer Slide Time: 05:40)



The slide is titled "Definitions of Learning" in yellow text on a dark blue background. It features a small gear icon in the top left corner. The content is organized into two numbered points, each followed by its source. The first point is a definition by Donald Michie (1991), and the second is a definition by Herbert Simon (1983). The slide also includes a standard presentation navigation bar at the bottom.

Definitions of Learning

1. A learning system uses sample data to generate an updated basis for improved [performance] on subsequent data from the same source and expresses the new basis in intelligible symbolic form.
Donald Michie (1991), Computer Journal
2. Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the same population more effectively the next time.
Herbert Simon (1983), Machine Learning 1

The first definition we will take up is the definition by Donald Michie in 1991. This definition states that a learning system uses sample data to generate an updated basis for improved performance on subsequent data from the same source and expresses the new basis in intelligible symbolic form. So we have a learning system which uses sample data which we call a training example. So there is sample data or there is some experience to go by and on the basis of this experience the system tries to generate a new model so that it can lead to improved performance on subsequent examples. So, the system uses input data to get a model which can help it to improve its performance on new data. But data that is coming from similar sort of examples data from the same source and this model

that has been learned is expressed in a symbolic form which can be understood and which can be manipulated. This is the definition by Donald Michie.

Let us look at the definition by another pioneer field Herbert Simon. this definition says, learning denotes changes in the system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the same population more effectively the next time. This is a similar definition as you can see. So this definition says, learning means change in the system and this change enables the system to perform better on similar tasks in future more effectively.

(Refer Slide Time: 08:00)



We see that they get several other definitions by several other practitioners in the field but we see that this type of learning mainly talk about there being some input data or experience in which some data system learns and the system tries to improve its performance. And this improvement in performance must be measurable in some way so there must be performance measure which must be improved due to learning. Therefore by learning from examples the system is able to improve its performance.

Secondly, some of these definitions also emphasize the comprehensibility. The new thing the system has learnt must be expressed in a form that will be understood.

(Refer Slide Time: 09:01)

Indian Institute of Technology, Kharagpur

Well-posed learning problem

- A computer program is said to learn
 - a Task T
 - given experience E
 - with respect to performance metric P

If its performance at T improves with E , as measured by metric P .

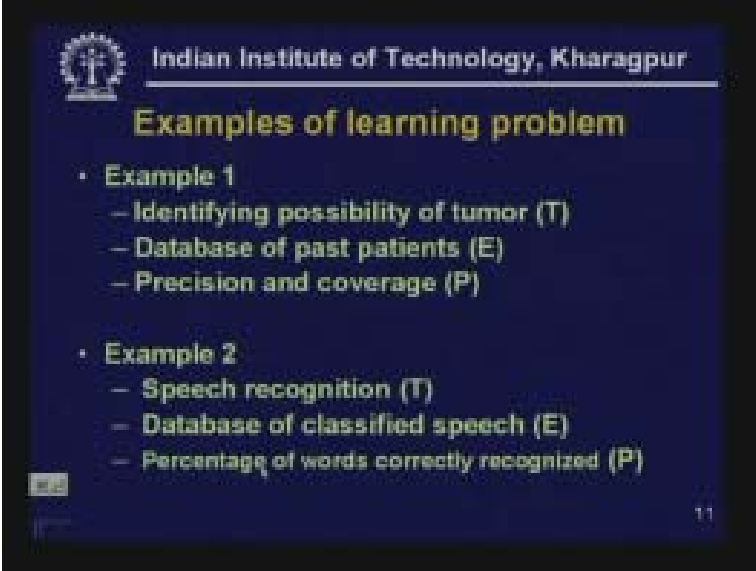
19

Therefore now let us simply define a well posed learning problem. A computer program is set to learn a task T , so T is the task the system is trying to learn. And what is the basis by which the system is learning this task T ?

The system is given some data which the experience e and the system is learning to improve its performance in task t with respect to its performance metric p . That is, the system's performance in the task T improves with experience e as measured by the performance metric p . Therefore this is the definition of learning we will accept. To put it in a more natural form, learning is the improvement of performance in some environment through the acquisition of knowledge resulting from experience in that environment. To understand this definition we need to go back to the framework of intelligent agent which we have been looking at in this course.

We have this structure, we have this agent and we have this environment. The agent takes action which changes the environment and the agent can sense the environment. The agent gets experience by interacting with the environment. And this results in acquisition of some knowledge and using this knowledge the agent can actually improve its performance in certain tasks. So this is the formal definition of learning. Now let us look at few examples of learning problem to make this more concrete.

(Refer Slide Time: 11:21)



The slide features the IIT Kharagpur logo and name at the top. The title 'Examples of learning problem' is in yellow. Below are two examples, each with a list of items. The first example is 'Identifying possibility of tumor (T)' with sub-points 'Database of past patients (E)' and 'Precision and coverage (P)'. The second example is 'Speech recognition (T)' with sub-points 'Database of classified speech (E)' and 'Percentage of words correctly recognized (P)'. A small 'CC BY' logo is in the bottom left, and the number '11' is in the bottom right.

Indian Institute of Technology, Kharagpur

Examples of learning problem

- Example 1
 - Identifying possibility of tumor (T)
 - Database of past patients (E)
 - Precision and coverage (P)
- Example 2
 - Speech recognition (T)
 - Database of classified speech (E)
 - Percentage of words correctly recognized (P)

CC BY 11

Let us say that we want to know whether a particular given patient has a possibility of having brain tumor. What is the experience?

We have the database of previous patients, the ones who have been diagnosed with tumor and ones who have been certified as having no tumor and we have accessed through the patient records which include different data about the patient as well as some images like MRI or something. So we have some data about past patients having tumor. And we wish to find out given a patient and the records we want to know whether this patient is likely to have tumor.

The second example is our task is to recognize speech. So the machine should be able to hear what we speak and it should be able to recognize the speech and take dictations. So in this case what can be used as an experience is a database of speech which has been already recognized and their transcripts are available. So we have speech and their transcripts that constitute the experience. And as a result of this experience the system should be able to learn and how can we measure whether the system is successful? We measure the percentage of correctly recognized words.

Therefore the success in the learning task can be measured by the accuracy or the precision. That is, number of examples which have been correctly labeled as correctly recognized. And the error can be measured by looking at the number of misclassifications, the number of words which have been wrongly recognized. So we want to recognize all correct cases. So, for tumor we want to label as positive all those patients who really have tumor and **none of the patients who do** not have tumor. We can measure it by accuracy or you can measure it by two terms precision and coverage. So precision accuracy is the number of examples that we label as having tumor and we want to know the percentage of them who actually have tumor and coverage means out of the patients who actually have tumor how many of them do we correctly recognize as

having tumor. These are the different measures we use for measuring the accuracy or the correctness of a learning task.

So the learner is in an environment learner is tempting to learn something about environment so that it can perform some of its tasks well. The learner is linked to some knowledge base from which it can draw and which stores the acquired knowledge. Therefore the learner has a knowledge base, there could be some prior knowledge in the knowledge base but as a result of learning the learner is able to update its knowledge base. And this knowledge is usually stored in some form of internal data structure.

You have studied logic and you have seen representational languages in logic, first order logic and you have also looked at certain other representation schemes and frames etc. Knowledge is basically stored in one of these representation schemes.

And what is experience?

Experience is basically derived from the perceptual input of the agent and the agent can take some action which is the output of the agent. And the performance of the agent is measured quantitatively by several aspects and depending on the task we decide what to use. Let us just review why study of machine learning is so important. There are many tasks that require an adaptive system that require a system which can learn.

(Refer Slide Time: 16:07)



For example, hand writing recognition, speech recognition are examples where an adaptive system is required. Learning is also useful as an alternative to hard coding a program. For example, suppose you want to develop a program which can play the game of chess. Now you could write a program in which for every different possible situation you can specify what move the agency. So you could hand code all the rules useful for playing chess. Therefore an alternative of writing such a program would be to provide the system with a database of chess, games and their outcomes and maybe the system can

apply the learning to learn to play a good game of chess **without even instructing the system to make the right move at the right situation**. But you give a large database of games by which the system can figure out. Providing database of games is usually easier than hand coding the rules. So you can save a lot of manual effort if a system is able to learn.

Also, the study of machine learning gives us an insight into human learning. For example, learning of language is a very non trivial task. So, trying to understand how we can make a machine learn a language can give us a clue as to how humans acquire language. Machine learning has been very useful in the curbing of data mining which helps systems to acquire hidden rules from data, which has opened a whole new area of applications. This is a new kind of capability that our systems are provided with. There are many types of learning and we can classify learning along different dimensions.

(Refer Slide Time: 18:39)



Supervised learning: In supervised learning the system is given labeled training examples. In these labeled training examples we have a set of examples and we also have the labels. So we have the inputs and the outputs of the examples. And unsupervised learning is the learning where there are no labels given. We only have the examples which are not pre-classified. So we have unclassified training example and there are situations where we like to learn from them. For these lectures we will be mainly concerned with supervised machine learning where we are given labeled examples. For the concept learning we have a set of labeled examples during training using which we will learn to classify unknown example. There are also other types of learning where we do not have labeled examples but we try to learn.

A third type of learning is reinforcement learning. In reinforcement learning we are not given just examples and each one of them is labeled but we are given a sequence of examples and at some points the system gets some reward or some punishments called

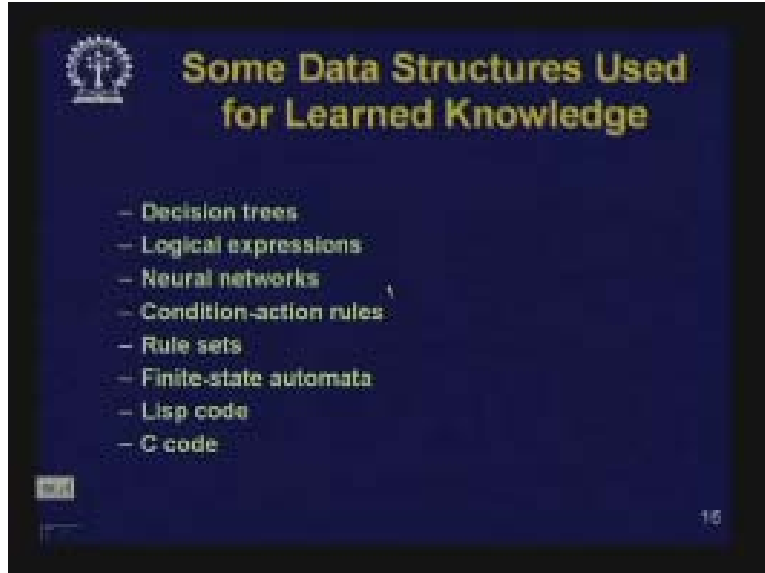
reinforcement. For example, when a system is playing a game of chess the system does not get to know whether each move is good or bad each move is not rated. But at the end of the game the system knows whether it has won or the game was a draw or it lost.

(Refer Slide Time: 21:30)



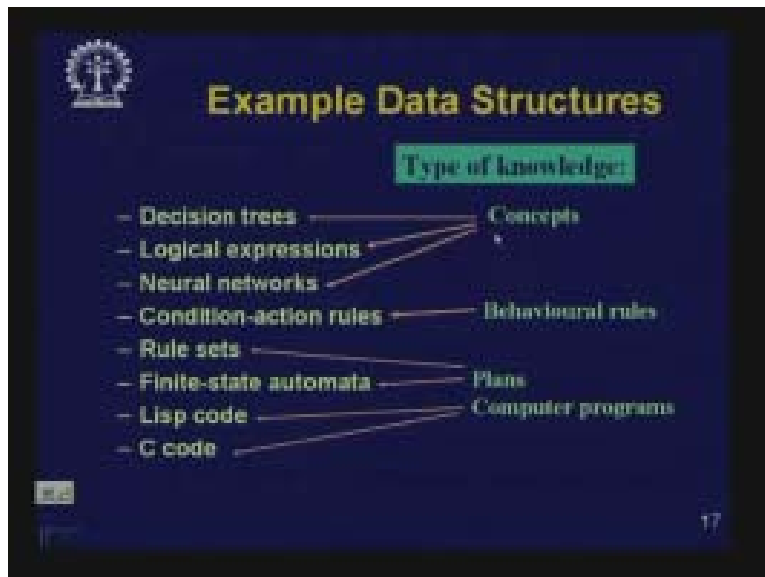
That is the reinforcement which is available at only certain points in the game. So, in reinforcement learning the system is trying to learn but it gets reinforcements only at a certain time. What are the different types of knowledge that can be acquired by learning? The types of knowledge can be declarative knowledge. Declarative knowledge can be expressed in terms of concepts, in terms of preferred value of parameters or in terms of grammar or in terms of taxonomies. The knowledge acquired by learning can alternatively be expressed as procedural knowledge. Procedural knowledge can be expressed in terms of rules, rule strengths, graphs or networks, computer programs and plans.

(Refer Slide Time: 21:38)



Certain data structures can be used for storing knowledge. Certain data structures are decision trees, logical expressions, neural networks, condition action rules, sets of rules, finite state automata and programs. For example, concepts when we want to represent concepts we can represent concepts by decision trees, or by logical expression or by neural networks. Behavioral rules can be expressed as condition action rules.

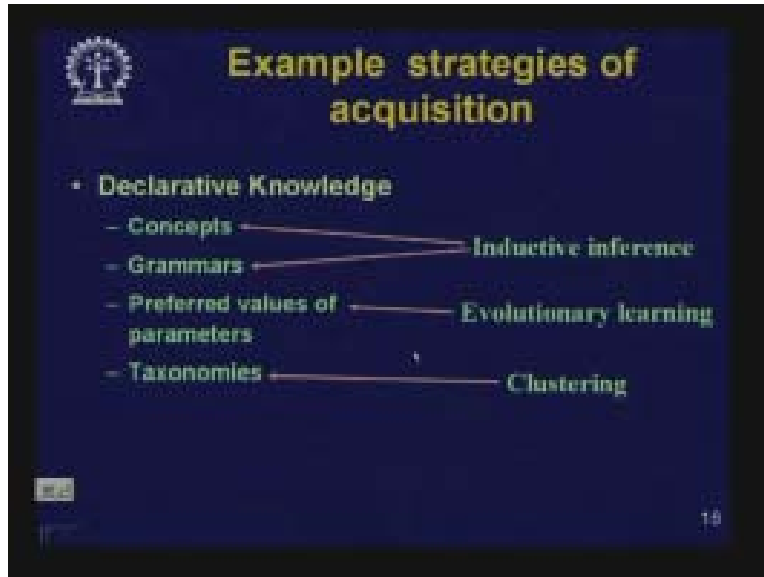
(Refer Slide Time: 22:06)



Plans can be expressed as sets of rules or by finite set automata. Computer program can be expressed as c code. And then there could be different strategies by which these

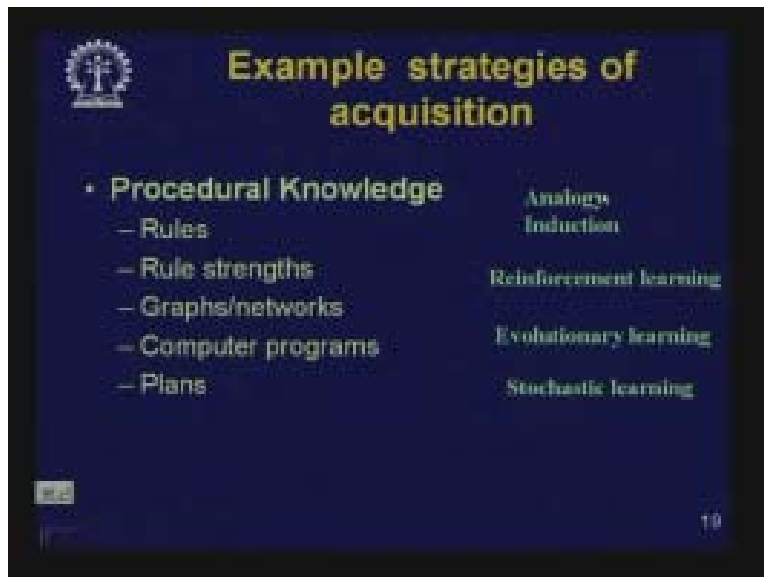
learning rules can be acquired. Inductive inference can be used to learn concepts and grammar.

(Refer Slide Time: 22:45)



Evolutionary learning or genetic algorithms can be used to learn certain preferred values of parameters. In unsupervised learning clustering can be used for learning taxonomies.

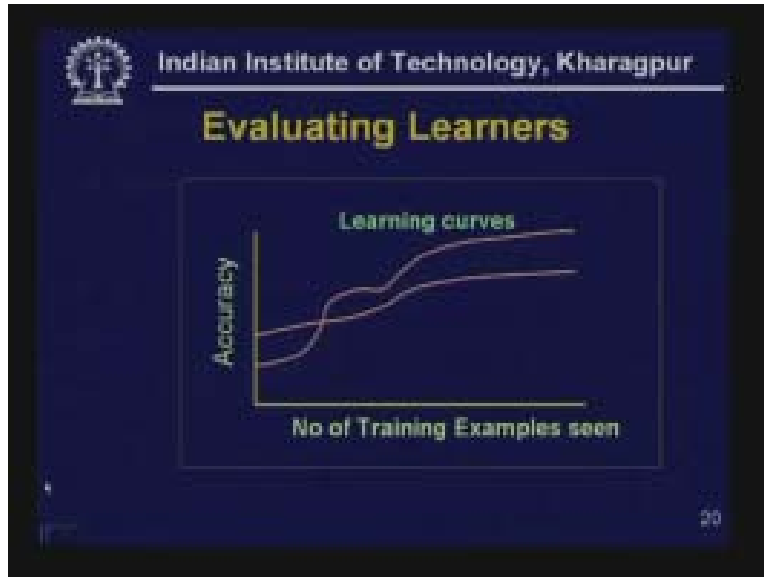
(Refer Slide Time: 23:02)



Then we can also have analogy or induction to learn rules, we can have reinforcement learning to learn plans or strategies or which you call policies, evolutionary learning,

stochastic learning etc. There is a wide choice of learning strategies and learning programs.

(Refer Slide Time: 23:31)



How do we evaluate learner?

Once we have a learning strategy typically what we do is we draw a learning curve. In a learning curve what we do is we plot the accuracy of the learner along the y axis, this axis is the accuracy and along the x axis we plot number for training example seen. So what we typically expect is that the accuracy of the learner will increase as the number of training examples increases. These are typical learning curves that we might get. So these learning curves plot the accuracy or precision in terms of the number of training example seen. Next we will talk about inductive learning for classification also known as concept learning. In fact in most of the lectures on learning in this course we will mainly be looking at the idea of concept learning.

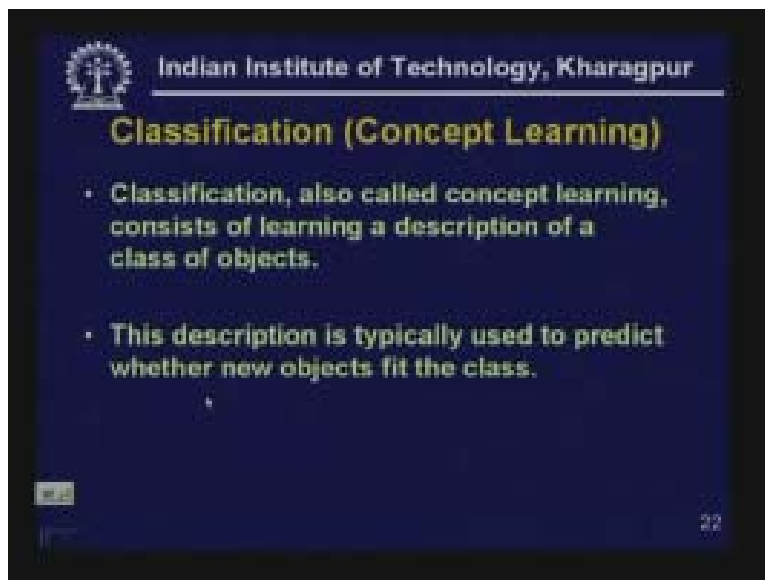
What is concept learning?

Concept learning or classification means learning at description of a class of objects. So we have some concept or some object whose description we wish to learn and concept learning is learning this description.

(Refer Slide Time: 24:26)



(Refer Slide Time: (24:44))

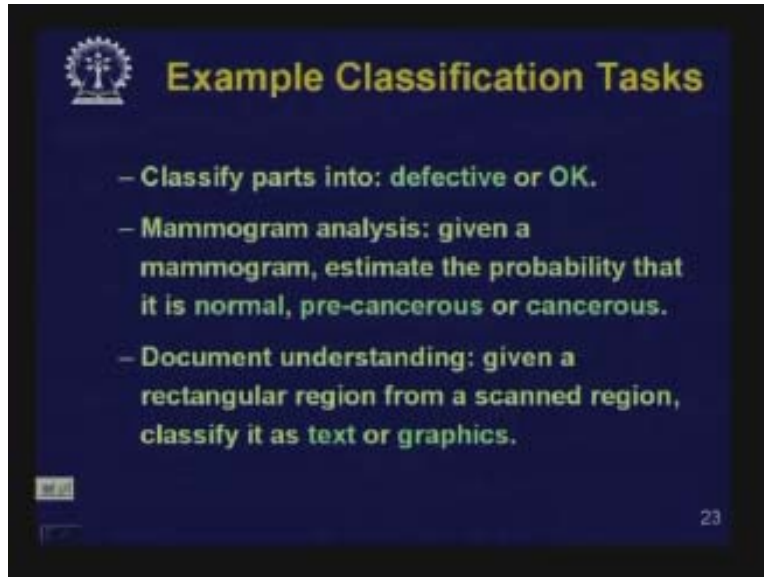


Why do we want to learn this description?

We want to use this description to predict the class of a new object. When we have a new example we want to know which class this new example belongs to so we try to learn the description of a class. So, if we are given an animal we want to know whether this animal can be classified as a tiger or not. And in the past we have seen several animals we are told they are tigers and we also have seen several animals for which we were told they are not tigers and from this we will form a model what is a tiger. And when given a new animal we will be able to know whether this animal is a tiger or not a tiger, so this is the essence of concept learning.

There are other examples of concept learning tasks some of them are described here. Suppose our objective is to classify parts as defective parts or ok parts. Second example is mammogram analysis. We are given a mammogram and we want to classify whether the mammogram is either normal or precancerous or cancerous.

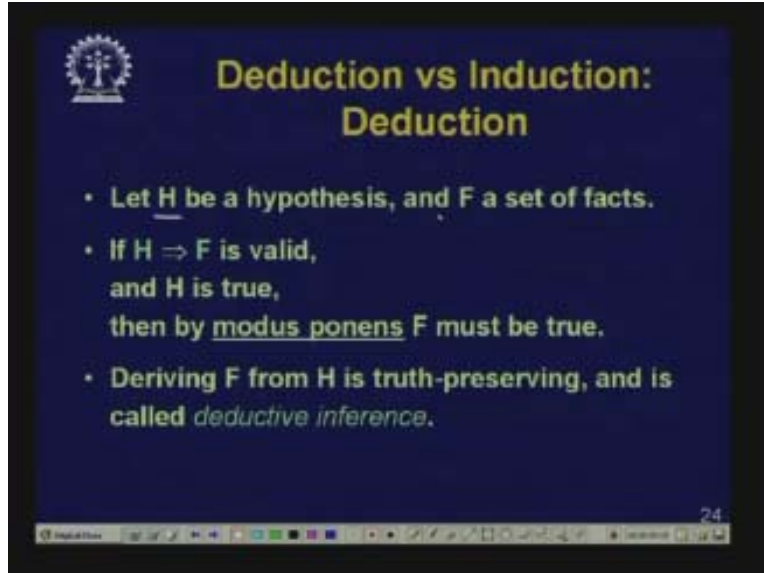
(Refer Slide Time: 26:35)



Example 3; in document understanding we are given a rectangular region from a scanned image and we should be able to say whether this is a text region or graphics region. So we want to recognize a text region from a graphics region. These are certain examples of concept learning tasks. Now we will be using inductive inference for concept learning. And let us see what we mean by inductive inference.

Suppose H is a hypothesis and F is a set of facts, suppose we know that H implies F, if H implies F is valid this rule implies f is valid and H is known to be true then by Modus ponens F must be true. H implies F is valid, the antecedent H is known to be true then logically by Modus ponens it follows that F must be true. This is an example of logical deduction. So we can derive F from H. this process is truth preserving and this is called deductive inference. Let us take an example. You know that all men are mortal and you know that Socrates is a man and you can conclude that Socrates is mortal. This is an example of deductive inference.

(Refer Slide Time: 27:09)



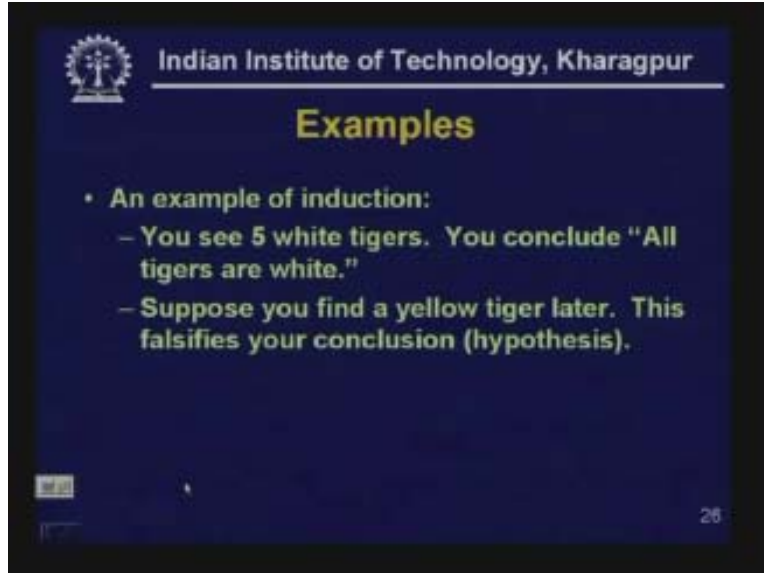
The slide features a dark blue background with a white logo in the top left corner. The title 'Deduction vs Induction: Deduction' is written in yellow. Below the title, there are three bullet points in white text. At the bottom right of the slide, the number '24' is visible.

- Let H be a hypothesis, and F a set of facts.
- If $H \Rightarrow F$ is valid, and H is true, then by modus ponens F must be true.
- Deriving F from H is truth-preserving, and is called *deductive inference*.

Next let us see what we mean by inductive inference. Suppose again let us assume that h is a hypothesis and f is a set of facts and again let us say that H implies F is known to be valid. Now suppose we know F , now knowing F in inductive inference we will try to derive H . Now from F H does not follow by deductive inference. But if F is false then H must be false, if F is true deductively we cannot say that H must be true but if F is false we can say that H must be false. So deriving H from F is falsity preserving but not truth preserving. So, if there are some facts which make f false then h must also be false. But in inductive inference we will derive H from F . Suppose for 10 days I have woken up and seen that in the morning the sky is blue and on the basis of this I form a hypothesis that in the morning the sky is always blue. This is an example of an inductive inference. On the basis of 10 data of the sky being blue in the morning I am inferring that the sky is always blue. This is an example of inductive inference or induction.

Another example of induction is; suppose you see five white tigers, you see five tigers and all of them are white and you conclude all tigers are white. This is an example of inductive inference.

(Refer Slide Time: 30:24)



Indian Institute of Technology, Kharagpur

Examples

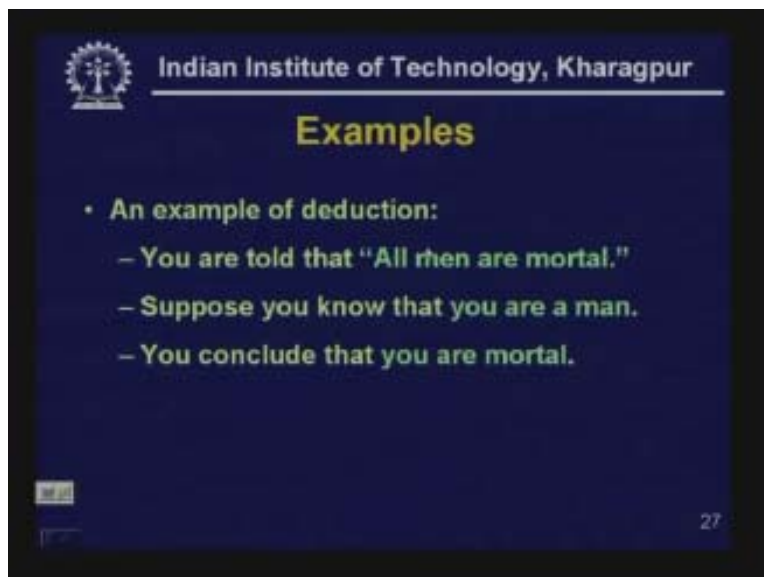
- An example of induction:
 - You see 5 white tigers. You conclude "All tigers are white."
 - Suppose you find a yellow tiger later. This falsifies your conclusion (hypothesis).

26

When can you be wrong?

Suppose now you find a yellow tiger now this falsifies your conjunction. So inductive inference is not a logical inference so it does not preserve the truth but it is a useful leave to take, it gives you new knowledge, knowledge that you have not already seen. An example of deduction which I already mentioned is suppose you told that all men are mortal and you know that you are man and you conclude that you are mortal. This is an example of deductive inference.

(Refer Slide Time: (31:10)



Indian Institute of Technology, Kharagpur

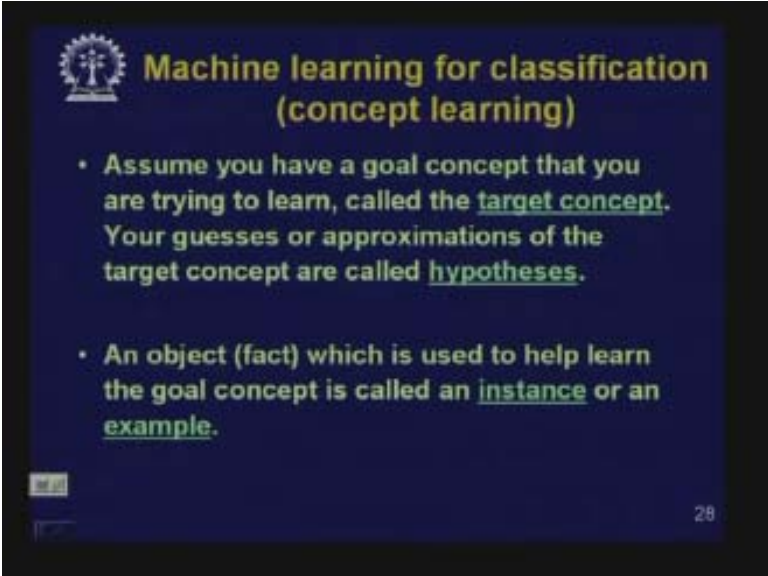
Examples

- An example of deduction:
 - You are told that "All men are mortal."
 - Suppose you know that you are a man.
 - You conclude that you are mortal.

27

What we saw earlier was an example of inductive inference. Now let us look at concept learning in more detail.

(Refer Slide Time: 31:30)



The slide is titled "Machine learning for classification (concept learning)" and features a logo in the top left corner. It contains two bullet points:

- Assume you have a goal concept that you are trying to learn, called the target concept. Your guesses or approximations of the target concept are called hypotheses.
- An object (fact) which is used to help learn the goal concept is called an instance or an example.

The slide number "28" is visible in the bottom right corner.

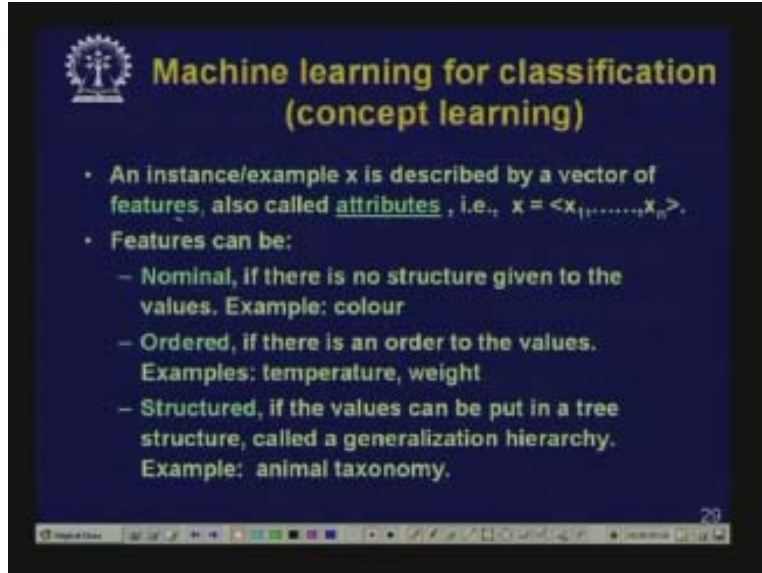
Suppose you have a goal concept that you are trying to learn so we will call this goal concept as a target concept. A target concept is a concept you are trying to learn. For example, you are trying to learn about a tiger so tiger is a target concept. And your guesses of the target concept is the hypotheses. So you are trying to learn the concept and you make hypothesis. You form a hypothesis about the description of the concept. So a hypothesis is your guess or approximation of the target concept.

How do you form hypothesis?

You form a hypothesis by looking at data, by looking at many instances of data or examples. So, an object which is used to help you learn the goal concept is called as instance or an example. An instance or an example is something which helps you learn the goal concept. Therefore typically an instance is described by a vector of features also called attributes.

Example of attributes; Suppose you are given a new animal and the attributes are color, number of legs, whether the animal has **viscous**, length of the body, whether the animal has fur and so on so these are the features. If you are given an animal whose color is white, it is furry, it has a tail, it is 6 ft long, it has **viscous** then this is a description of an instance of an animal. Therefore an instance is described by vector of features, features are also called attributes. So, for example the instance is given by x_1, x_2, x_n and x_1 is the value of the first feature, x_n is the value of the nth feature. There are different types of features that we can have, features can be nominal.

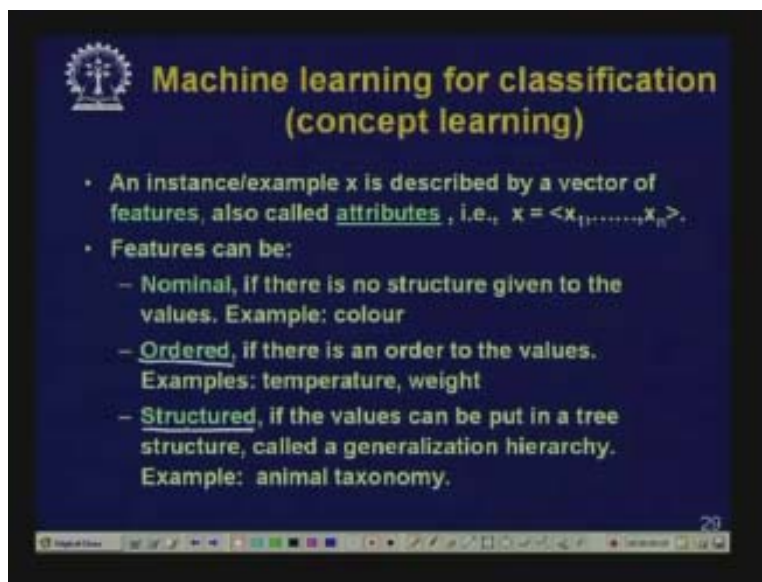
(Refer Slide Time: 32:45)



Machine learning for classification (concept learning)

- An instance/example x is described by a vector of features, also called attributes, i.e., $x = \langle x_1, \dots, x_n \rangle$.
- Features can be:
 - Nominal, if there is no structure given to the values. Example: colour
 - Ordered, if there is an order to the values. Examples: temperature, weight
 - Structured, if the values can be put in a tree structure, called a generalization hierarchy. Example: animal taxonomy.

(Refer Slide Time: 34:53)



Machine learning for classification (concept learning)

- An instance/example x is described by a vector of features, also called attributes, i.e., $x = \langle x_1, \dots, x_n \rangle$.
- Features can be:
 - Nominal, if there is no structure given to the values. Example: colour
 - Ordered, if there is an order to the values. Examples: temperature, weight
 - Structured, if the values can be put in a tree structure, called a generalization hierarchy. Example: animal taxonomy.

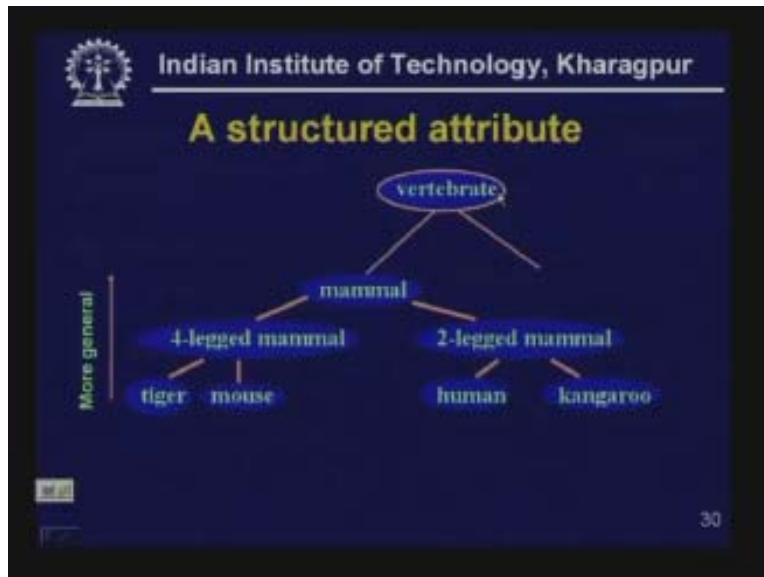
For example, color of the animal, color can be red, blue, yellow, green so these are some specific values they are not directly linked to each other. Therefore color is a nominal attribute similarly suppose you are trying to learn an object whether the object is a table and one of the attributes it can have is from what material the object is made of? And the material can be wood, steel, glass etc. So these are nominal attributes.

Secondly, you can have numeric attributes. Usually the attributes can be ordered the other attributes are also in an order. Order means there is an ordering to the values. For example, length of the animal is the order attribute; it is 5 ft, 6 ft, 4 ft, 8 ft so they have

some relationship within themselves. This is an order attribute. Temperature, weight, length are examples of order attributes.

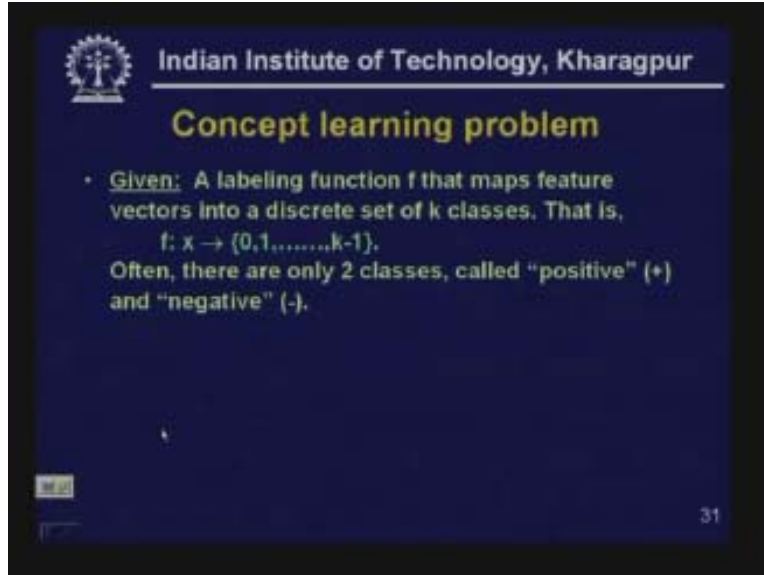
Thirdly attributes can be structured. There is some structure in the order of the values but they are not fully ordered. For example, if the values can be put into some sort of generalization hierarchy or some partial order. For example, if you consider the animal taxonomy, this is an example of animal taxonomy.

(Refer Slide Time: 35:14)



Suppose you have the class vertebrates and you know that animals can be classified into invertebrates and among vertebrates you can have mammals, reptiles etc and among mammals birds, fish are vertebrates, so among mammals you can classify them as four legged mammals and two legged mammal, under four legged mammal we have tiger, mouse, deer etc and under two legged mammal we have humans, kangaroo etc. Therefore this is an example of classification hierarchy and this is also an example of a structured attribute and in this direction we have more general classes.

(Refer Slide Time: 36:01)



The slide features the IIT Kharagpur logo and name at the top. The title 'Concept learning problem' is in yellow. The main text is in white on a dark blue background. It defines a labeling function f that maps feature vectors to a discrete set of k classes, with the mathematical expression $f: x \rightarrow \{0, 1, \dots, k-1\}$. It also notes that often there are only two classes, 'positive' (+) and 'negative' (-). A small navigation icon is in the bottom left, and the number '31' is in the bottom right.

So the concept learning problem formally can be described as follows: We are given labeling function f , f is a underlying function for describing the concept. This function f maps feature vectors into some classes. Therefore we have a discrete set of k classes that are k possible classes.

For example, we have two classes tiger or not tiger or we can have three classes whether it is a tiger or a lion or a deer. So, in general we can have a finite set of classes so we have k classes and we have an actual underlying function which maps to instance to one of these classes. So the function f and maps an input instances to one of these classes. In a special case we may have only two classes and in such a case we may say that one of the classes is positive and the other one is negative. We are given some training examples.

(Refer Slide Time: 37:13)

Indian Institute of Technology, Kharagpur

Concept learning problem

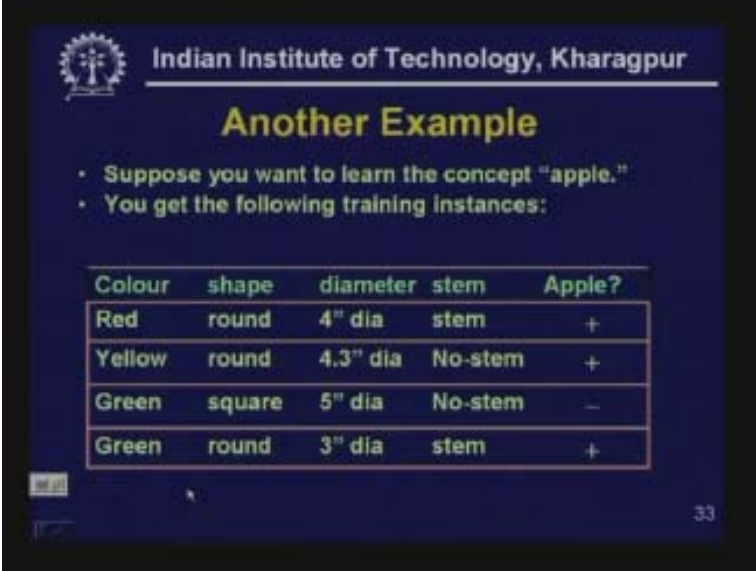
- **Given:** A labeling function f that maps feature vectors into a discrete set of k classes. That is, $f: x \rightarrow \{0, 1, \dots, k-1\}$. Often, there are only 2 classes, called "positive" (+) and "negative" (-).
- Represent each **training example** as a pair $(x, f(x))$. These are the examples that will be used for learning the concept.
- **Problem:** From a set of $(x, f(x))$ pairs, learn the target concept f .

31

Each training example is a pair, the instance and its classification. So each training example is given as a pair the instance and its class and we have a set of such training examples. So using these training examples we want to learn an approximation of f . Therefore from the set of x $f(x)$ pair we have got we want to learn the target concept f . Therefore f is what we wish to learn. So given a set of x of $f(x)$ pairs we want to infer f we want to inductively infer f .

Now, if you are given a finite sample and you are not shown all the instances it is not really possible to guess the correct value of f with absolute certainty. So we will apply inductive inference and we will try to find some pattern in the training examples and we will assume that this pattern will hold for future examples also. This is an example of training set, we have an instance x is equal to 1 then the classification is 1, x is equal to 2 $f(x)$ is equal to 3 $f(x)$ is equal to 9 x is equal to 4, $f(x)$ is equal to 16 and we want to know if x is 5 then what is $f(x)$? This is an example of a training set that we are give. And if you want to learn the function f in this case one good guess is that the function f is a square function that is $f(x)$ is x square but this not really a discrete learning problem but this is an example of learning a function.

(Refer Slide Time: 39:19)



Indian Institute of Technology, Kharagpur

Another Example

- Suppose you want to learn the concept "apple."
- You get the following training instances:

Colour	shape	diameter	stem	Apple?
Red	round	4" dia	stem	+
Yellow	round	4.3" dia	No-stem	+
Green	square	5" dia	No-stem	-
Green	round	3" dia	stem	+

33

Another example; suppose you want to learn the concept of an apple whether an object is an apple or not, you are given some training examples. For training examples these are the features that are given. For every object that you are given you know the color, shape, diameter, whether it has a stem and then it is labeled whether the object is an apple or not. You have got four examples.

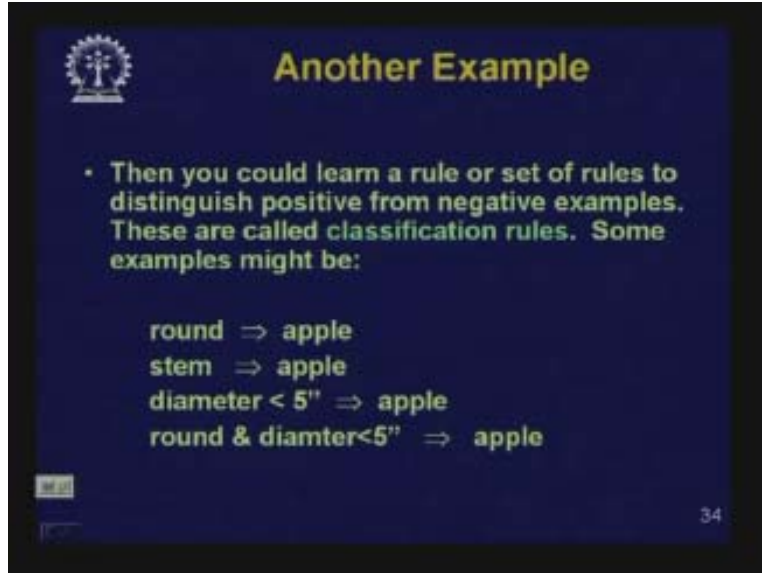
The first example the color is red, shape is round, it has a diameter of four inch, it has a stem and you know it is an apple.

The second example is, the color is yellow, shape is round, diameter is 4.3 inch. It does not have a stem it is an apple.

The third example; the color is green, shape is square, diameter is 5 inch, no stem, not an apple, this is a negative example of an apple.

Fourth example; the color is green, shape is round, diameter is 3 inch, it has a stem and it is an apple. Now you could learn a rule or a set of rules to distinguish positive examples of apple from negative examples. These rules are called classification rules. For example, the rules could be, round means apple or stem implies apple or diameter less than 5 inch means apple or round and diameter less than 5 inch means apple. So these are some possible rules you can hypothesize about an apple.

(Refer Slide Time: 40:45)

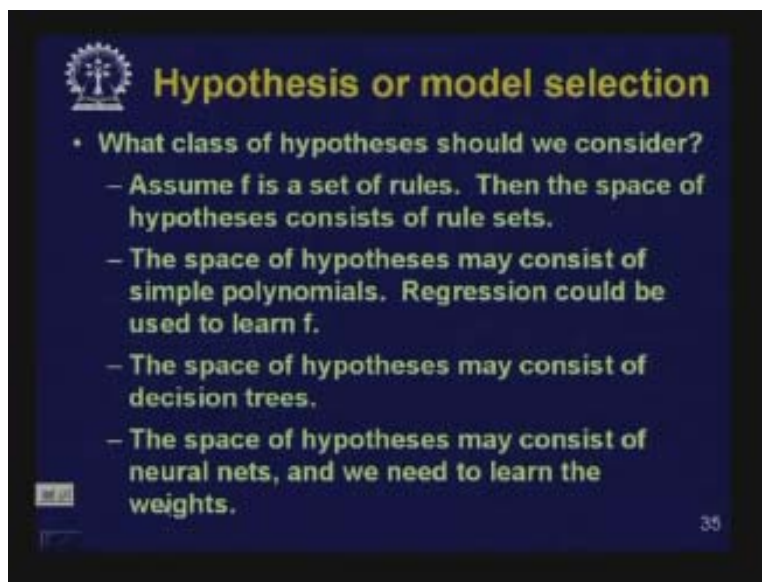


The slide features a dark blue background with a white logo in the top left corner. The title "Another Example" is written in yellow at the top center. Below the title, a bullet point in white text reads: "Then you could learn a rule or set of rules to distinguish positive from negative examples. These are called classification rules. Some examples might be:". Underneath, four rules are listed in white text: "round \Rightarrow apple", "stem \Rightarrow apple", "diameter < 5'' \Rightarrow apple", and "round & diamter<5'' \Rightarrow apple". A small white icon is in the bottom left, and the number "34" is in the bottom right.

Another Example

- Then you could learn a rule or set of rules to distinguish positive from negative examples. These are called classification rules. Some examples might be:
 - round \Rightarrow apple
 - stem \Rightarrow apple
 - diameter < 5'' \Rightarrow apple
 - round & diamter<5'' \Rightarrow apple

(Refer Slide Time: 41:07)



The slide features a dark blue background with a white logo in the top left corner. The title "Hypothesis or model selection" is written in yellow at the top center. Below the title, a bullet point in white text reads: "What class of hypotheses should we consider?". This is followed by four sub-bullets in white text: "Assume f is a set of rules. Then the space of hypotheses consists of rule sets.", "The space of hypotheses may consist of simple polynomials. Regression could be used to learn f.", "The space of hypotheses may consist of decision trees.", and "The space of hypotheses may consist of neural nets, and we need to learn the weights." A small white icon is in the bottom left, and the number "35" is in the bottom right.

Hypothesis or model selection

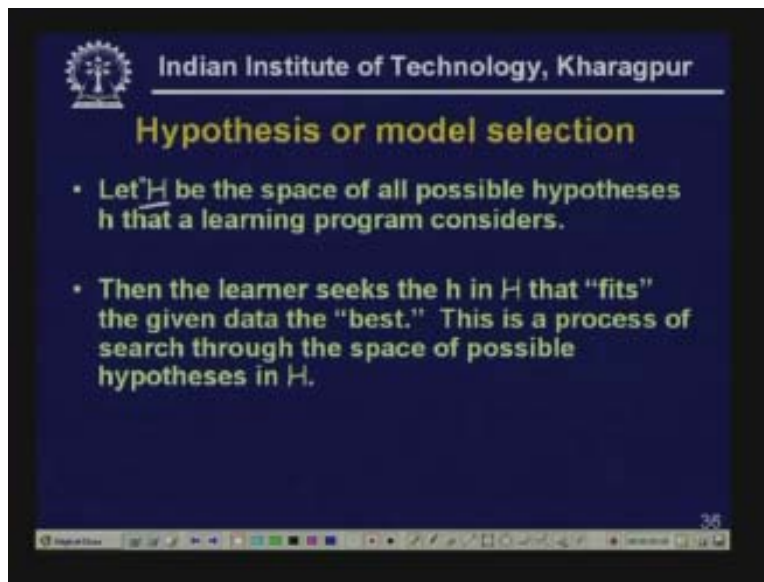
- What class of hypotheses should we consider?
 - Assume f is a set of rules. Then the space of hypotheses consists of rule sets.
 - The space of hypotheses may consist of simple polynomials. Regression could be used to learn f.
 - The space of hypotheses may consist of decision trees.
 - The space of hypotheses may consist of neural nets, and we need to learn the weights.

In this case we are looking at rules expressed in a particular type of language. These rules are given in terms of the features. But these features are expressed in some language. For example, round and diameter is less than five inch so here we are expressing the antecedent as a conjunction of constraints and features. Similarly we could express rule as a disjunction. We could say round or diameter less than 5 inch implies apple. So this is the example of disjunctive rule. Or we could express it other forms. We could decide a language to express the rules. And this defines the set of hypothesis that we are considering.

What type of hypothesis should we consider?

Suppose f is the set of rules then the space of hypothesis is all rule sets. The space of hypothesis could be simple polynomials; the space of hypothesis could be decision trees, the space of hypothesis we are considering could be neural networks. Therefore we have a choice about what type of hypothesis we are going to consider. We are trying to learn a function f which we do not know and we are trying to find an approximation to f , we are trying to make a hypothesis. We will have to choose this hypothesis from a set of hypothesis that we will consider. This is called our hypothesis phase.

(Refer Slide Time: 43:25)



Indian Institute of Technology, Kharagpur

Hypothesis or model selection

- Let H be the space of all possible hypotheses h that a learning program considers.
- Then the learner seeks the h in H that "fits" the given data the "best." This is a process of search through the space of possible hypotheses in H .

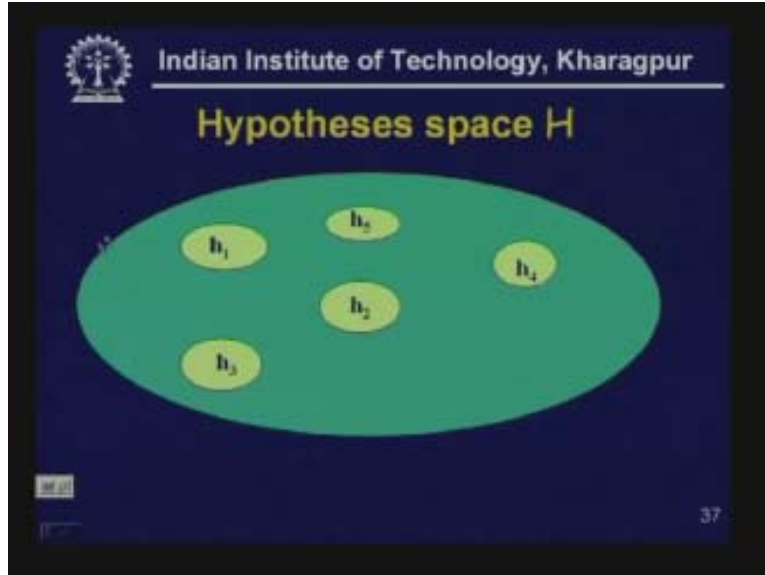
And what set of hypothesis we should consider?

This is what we must decide. So we will express by H the space of all possible hypothesis that a learning program consider. Then our objective would be to find one hypothesis which is a member of this hypothesis space. We want to find that hypothesis which is the best. We want to find the best hypothesis of this hypothesis space.

What is the best hypothesis?

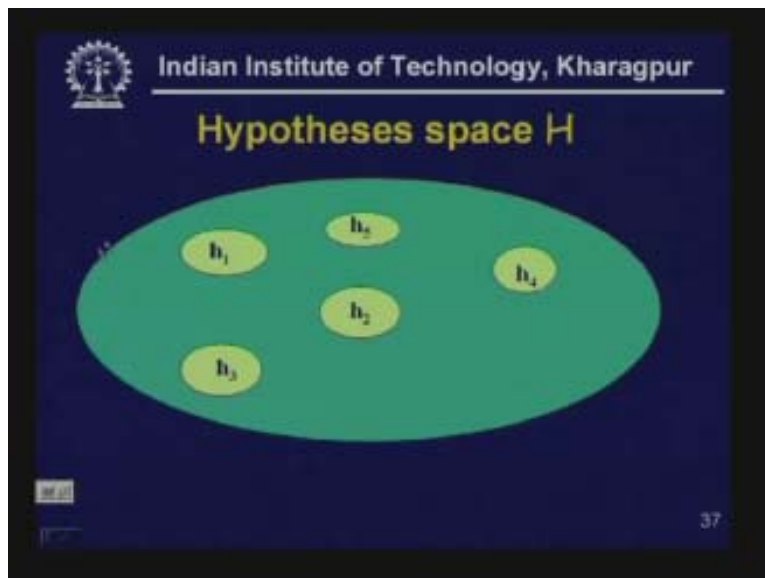
It is that which fits the given data the best. Therefore this is a process of search. In order to find its best hypothesis we will search through this hypothesis space to find that hypothesis which fits the data in the best possible manner.

(Refer Slide Time: 44:14)



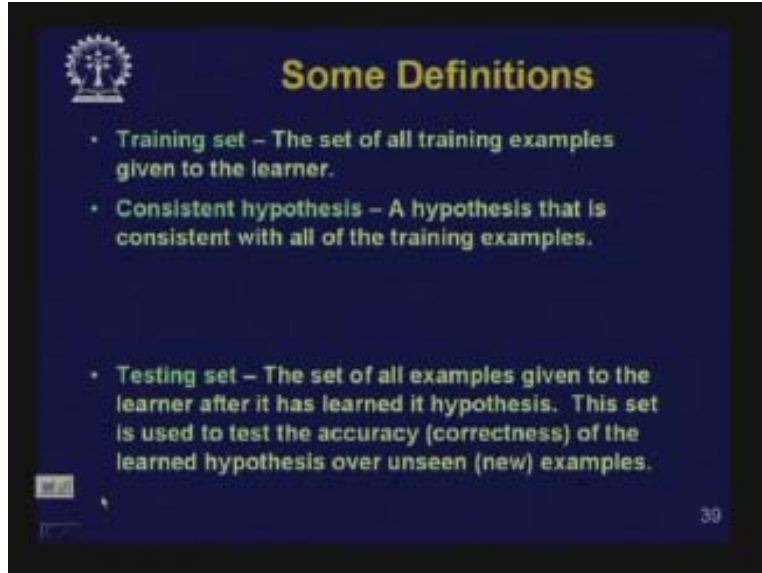
This is an example of a hypothesis space. In this hypothesis space there is a large number of hypothesis. For example h_1 h_2 h_3 h_4 h_5 are some of the hypothesis in this hypothesis space and we want to select the best of the hypothesis in this hypothesis space.

(Refer Slide Time: 44:14)



Suppose you want to learn the concept **apple** and this is an example of a hypothesis space, you have a hypothesis here like; round and diameter less than 5 degree, red and round has stem, not a square, round etc are some of the hypothesis that are present in the hypothesis space and we want to select the best of these hypothesis.

(Refer Slide Time: 45:10)



Some Definitions

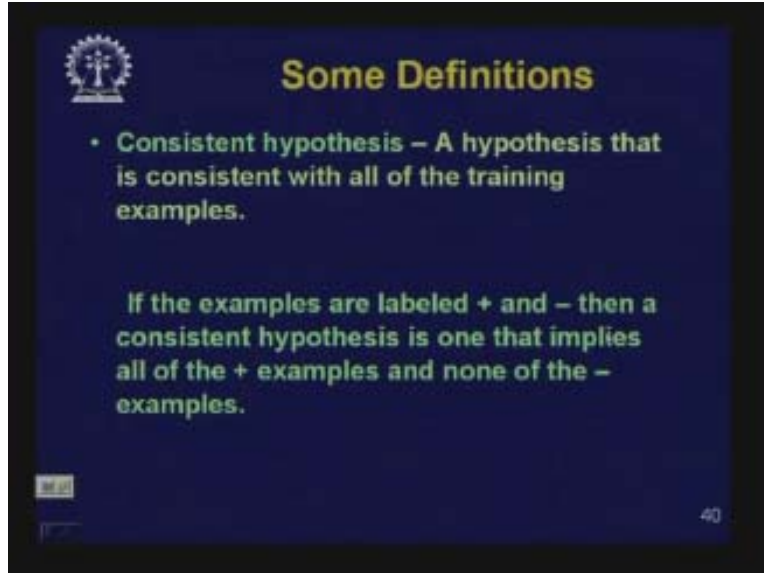
- **Training set** – The set of all training examples given to the learner.
- **Consistent hypothesis** – A hypothesis that is consistent with all of the training examples.
- **Testing set** – The set of all examples given to the learner after it has learned its hypothesis. This set is used to test the accuracy (correctness) of the learned hypothesis over unseen (new) examples.

30

Some definitions: A training set is the set of all examples that are given to the learner. A testing set is a set of examples on which the learner tests his hypothesis. So, on the basis of training set the learner will learn the hypothesis and the learner will try to evaluate how good the hypothesis is by looking at another set of examples which is called the test set. So the learner, on looking at the training set will find a hypothesis. The hypothesis the learner has found will be set to be consistent if it is consistent with all the training examples. That is, the hypothesis must predict the correct label of all the training examples and such a hypothesis is called consistent hypothesis. We cannot guarantee that a consistent hypothesis will necessarily always give the correct labeling for each test example because we have not seen the test example but we can check whether the hypothesis correctly labels the training example.

A testing set is the set of all examples that are given to the learner after it has learnt the hypothesis on which its accuracy will be tested.

(Refer Slide Time: 47:04)



Some Definitions

- **Consistent hypothesis** – A hypothesis that is consistent with all of the training examples.

If the examples are labeled + and – then a consistent hypothesis is one that implies all of the + examples and none of the – examples.

40

Suppose the examples are labeled as plus and minus then the consistent hypothesis is a hypothesis that implies that labels are positive for all the plus examples and none of the minus examples, that is a consistent hypothesis.

(Refer Slide Time: 47:21)



Indian Institute of Technology, Kharagpur

False positive and false negative

- An example is a **false negative** for a hypothesis if the hypothesis says it should be negative but it's positive.
- An example is a **false positive** for a hypothesis if the hypothesis says it should be positive but it's negative.

Instance space

True concept

hypothesis

41

Now let us look at this diagram. Suppose this one is the instance space which consists of all possible examples and suppose this is the true function f that our system is trying to learn. So this f labels the instances inside the circle as positive and the instances outside the circle as negative. Now this other circle is the hypothesis h that your learner finds. In this region the h and f agree the label of instances. The h makes mistakes in this blue

region and in this black region. So these are the two regions in which h makes a mistake. An example is called to be a false, negative if the hypothesis says it should be negative but it is actually positive. In this blue region your hypothesis states that these examples must be negative but these examples are actually positive according to f so these are falsely negative by your hypothesis. So this is a zone of error of your hypothesis. In this green region your hypothesis says that this instance should be positive but they are actually negative according to f .

This is another error zone and this is said to be the false positive for your hypothesis. So an example is a false negative for a hypothesis if the hypothesis says it would be negative but it is actually positive. An example is false positive for hypothesis if the hypothesis says it should be positive but it is negative. This big thing is the instance space, all the instances are here and this is a particular concept, inside the circle we have the positive example, outside the circle we have negative examples, and this other circle is your hypothesis. Inductive bias is another definition.

(Refer Slide Time: 50:21)



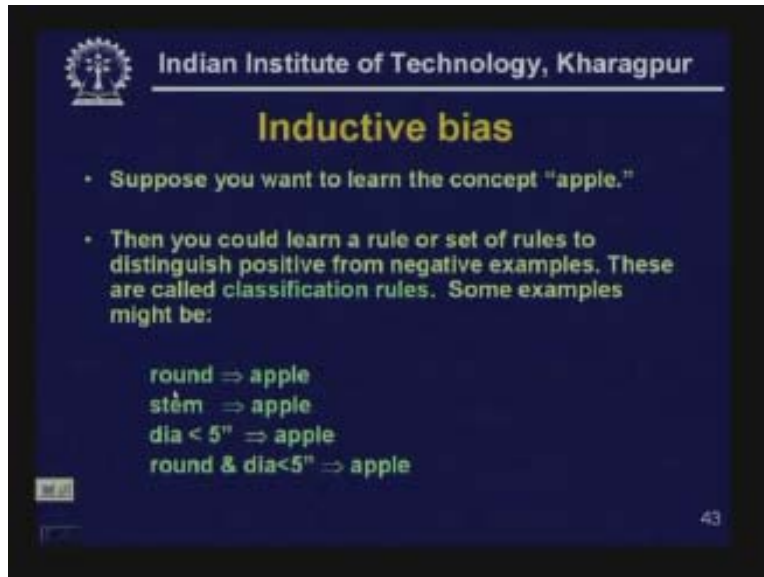
When you are making an inductive inference on the basis of looking at some examples you are making a hypothesis. That hypothesis may not be fully correct. For example, if you are making a jump, you are making a conceptual jump, you are assuming something extreme so when you see some data you could infer different possible hypothesis.

Which one of them would you select?

You could try to find out that hypothesis which is consistent but suppose you find several hypothesis which you are consistent with your training set, like in the apple example we could find several hypothesis which are consistent the four examples so which one of them would you select? In order to select one of them we use the concept of inductive bias. Therefore inductive bias is a bias or a preference for one hypothesis over another. And there are different types of inductive bias which are used by different learners. For

example, simplicity could be a bias or most general hypothesis could be a bias or choose the most specific hypothesis could be a bias. Therefore among the hypothesis which is both consistent we choose one of them according to some bias we have.

(Refer Slide Time: 51:59)



Indian Institute of Technology, Kharagpur

Inductive bias

- Suppose you want to learn the concept "apple."
- Then you could learn a rule or set of rules to distinguish positive from negative examples. These are called classification rules. Some examples might be:

```
round => apple
stem => apple
dia < 5" => apple
round & dia < 5" => apple
```

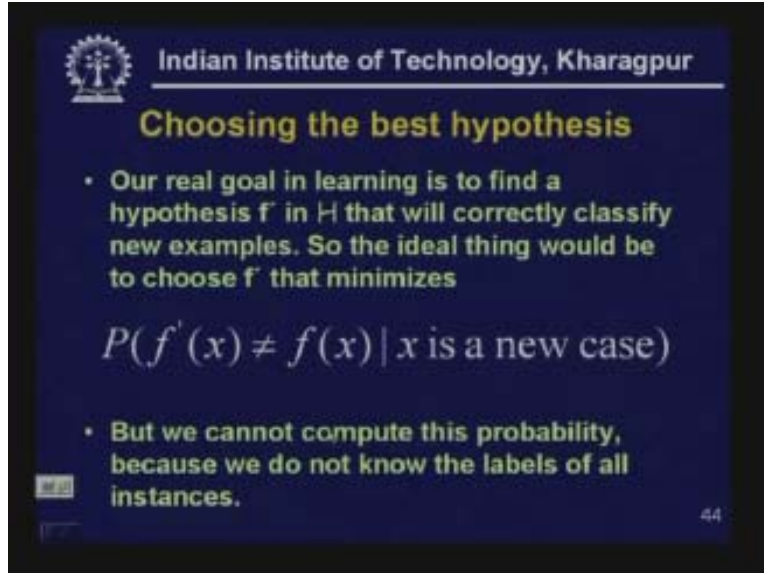
43

For example, suppose you want to learn the concept apple then you could learn a set of rules to distinguish positive from negative examples. These are called classification rules and these are rules we have already seen. Then you could say in order to select among rules you apply a bias this is called inductive bias.

How do you choose the best hypothesis?

One way in which you could choose the best hypothesis is, among two hypotheses which are equally consistent on the training examples you should choose a hypothesis that is likely to agree with new examples. So our real goal in machine learning is to find a hypothesis f' in h such that it will correctly classify new examples. So the ideal thing would be to choose that f' so that probability that $f'x$ does not agree with $f(x)$ is smallest.

(Refer Slide Time: 52:55)



Indian Institute of Technology, Kharagpur

Choosing the best hypothesis

- Our real goal in learning is to find a hypothesis f' in H that will correctly classify new examples. So the ideal thing would be to choose f' that minimizes

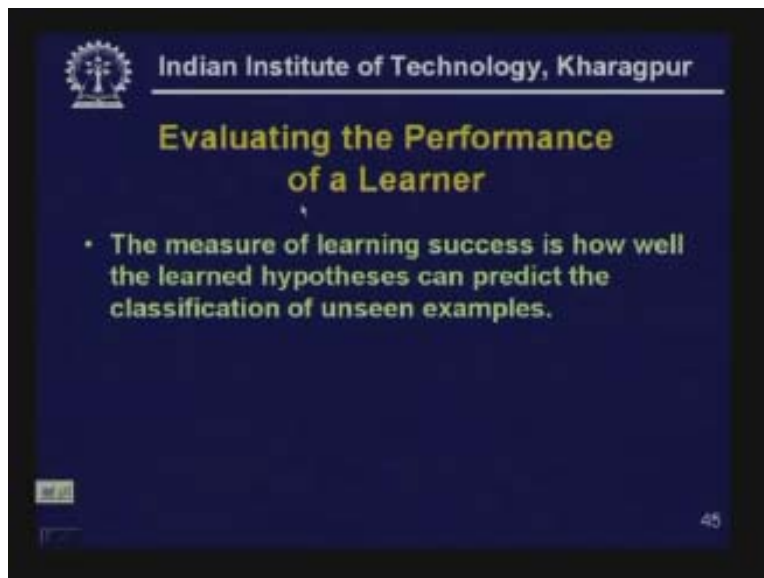
$$P(f'(x) \neq f(x) \mid x \text{ is a new case})$$

- But we cannot compute this probability, because we do not know the labels of all instances.

44

We want to choose that f' so that it agrees with f most of the time. But we cannot compute this probability because we do not before hand know all the instances all the labels of instances etc.

(Refer Slide Time: 53:35)



Indian Institute of Technology, Kharagpur

Evaluating the Performance of a Learner

- The measure of learning success is how well the learned hypotheses can predict the classification of unseen examples.

45

How we will evaluate the performance of the learner?

We will evaluate the performance of the learner by seeing how well the learned hypothesis can predict the classification of unseen examples, the examples that we have not yet seen.

(Refer Slide Time: 53:53)

Machine learning programs for classification (concept learning)

- **Problem:** From a set of training examples: $D = \langle x, f(x) \rangle$ pairs, learn the target concept f .
- **Method for evaluation:** Measure the predictive performance of the learned hypothesis on a set of test examples (called the test set).
- The training and test data must be disjoint.

48

If we use the training set for classification we will not really understand whether hypothesis is performing well on unknown examples. So we should apply the hypothesis we learnt from the training set and a new data. So, test set should be **disjoint** with the training set. So, in a concept learning our problem is to learn the target concept f from a set of training examples. We evaluate we measure the predictive performance of the learned hypothesis on a set of test examples called the test set. And we require that the training and test set are different from each other so that there is no bias.

(Refer Slide Time: 54:54)

Indian Institute of Technology, Kharagpur

Questions

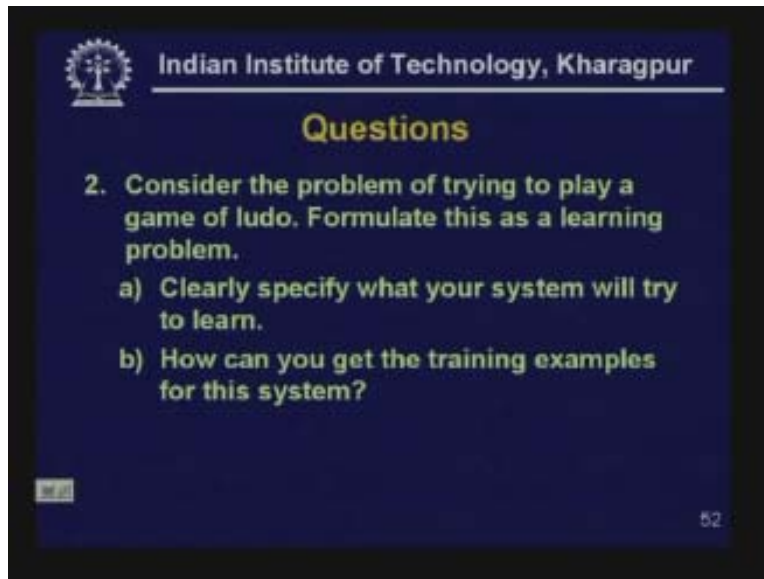
1. Consider the problem of trying to recognize handwritten digits. Formulate this as a concept learning problem.
 - a) Clearly specify what can be the possible features.
 - b) How do you get the training set and the test set?
 - c) How will you measure the performance of your learning algorithm?

51

Few questions to consider:

- 1) Consider the problem of trying to recognize hand written digits. You have to formulate this as a concept learning problem. Specifically you have three tasks; a) clearly specify what are the possible features?
 - b) How do you get the training set and the test set?
 - c) How will you measure the performance of your learning algorithm?

(Refer Slide Time: 55:36)



Indian Institute of Technology, Kharagpur

Questions

2. Consider the problem of trying to play a game of ludo. Formulate this as a learning problem.
 - a) Clearly specify what your system will try to learn.
 - b) How can you get the training examples for this system?

52

- 2) Consider the problem of trying to play a game of Ludo, formulate this as a learning problem. Clearly specify what your system will try to learn.
 - b) How can you get the training examples for this system?