SNS COLLEGE OF TECHNOLOGY

Coimbatore-35

An Autonomous Institution

Accredited by NBA – AICTE and Accredited by NAAC – UGC with 'A+' Grade

Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai

DEPARTMENT OF AI&ML

FOUNDATIONS OF ARTIFICIAL INTELLIGENCE

II YEAR - III SEM

UNIT 5 – **Decission Tress**

Decission Tree

- Learning from Observation
- Inductive Learning
- Decision Trees
- Explanation Based Learning
- Statistical Learning Methods
- Reinforcement Learning

Decission Tree

Decision Tree

- A decision tree takes as input an object or situation described by a set of attributes and returns a "decision"-the predicted output value for the input.
- The input attributes can be discrete or continuous.
- we assume discrete inputs, Then output value can also be discrete or continuous;
- Learning a discrete-valued function is called classification learning;
- Learning a continuous function is called regression learning.
- We will concentrate on *Boolean* classification, wherein each example is classified as true (positive) or false (negative).

Example – Restaurant – wait for a table

- The aim is to learn the **goal predicate** WillWait.
- The attributes:
  - 1. Alternate: whether there is a suitable alternative restaurant nearby.
  - 2. Bar: whether the restaurant has a comfortable bar area to wait in.
  - 3. Fri/Sat: true on Fridays and Saturdays.
  - 4. *Hungry:* whether we are hungry.
  - 5. Patrons: how many people are in the restaurant (values are None, Some, and Full).
  - 6. Price: the restaurant's price range ($, $$, $$$).
  - 7. Raining: whether it is raining outside.
  - 8. Reservation: whether we made a reservation.
  - 9. Type: the kind of restaurant (French, Italian, Thai, or burger).
  - 10. WaitEstimate: the wait estimated by the host (0-10 minutes, 10-30, 30-60, >60).

Example

Example – Restaurant – wait for a table...

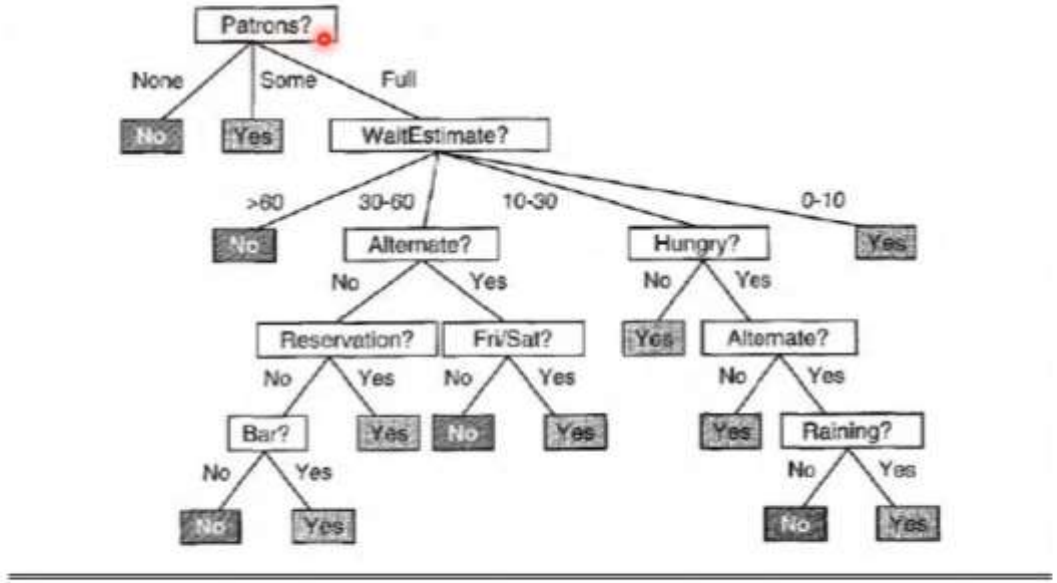
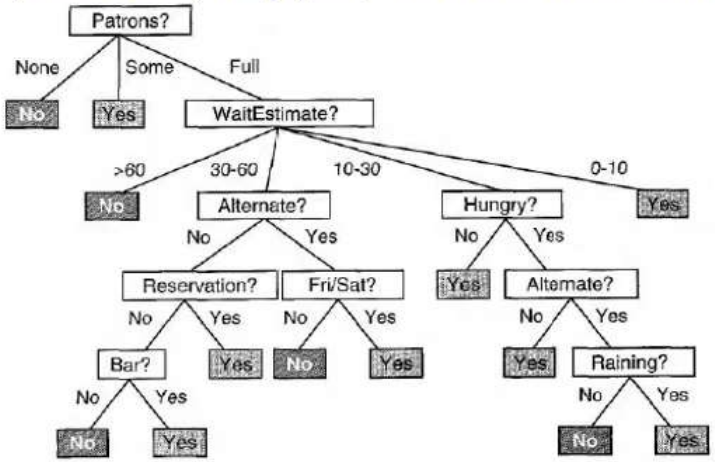


Figure 18.2 A decision tree for deciding whether to wait for a table.

- Notice that the tree does not use the Price and Type attributes, considering them to be irrelevant.
- Examples are processed by the tree starting at the root and following the appropriate branch until a leaf is reached.
- An example with Patrons = Full and WaitEstimate = 0-10 will be classified as positive (i.e., yes, we will wait for a table).



Choosing Attributes Test

Choosing Attribute Tests

- The scheme used in decision tree learning for selecting attributes is designed to minimize the depth of the final tree.
- The idea is to pick the attribute that goes towards an exact classification of the examples.
- A perfect attribute divides the examples into sets that are all positive or all negative.
- The Patrons attribute is not perfect, but it is fairly good.

Assesing the performance

Assessing the performance of the learning algorithm

- A learning algorithm is good if it produces hypotheses, when it predict the classifications of unseen examples.
- we can assess the quality of a hypothesis by checking its predictions against the correct classification, is called as Test Set
- Then the learning algorithm will perform the following methodology
  - 1. Collect a large set of examples.
  - 2. Divide it into two disjoint sets: the **training set** and the **test set.**
  - **3.** Apply the learning algorithm to the training set, generating a hypothesis h.
  - 4. Measure the percentage of examples in the test set that are correctly classified by h.
  - 5. Repeat steps 2 to 4 for different sizes of training sets and different randomly selected training sets of each size.

Contd...

Assessing the performance of the learning algorithm...

- The result of this procedure is a set of data, that can be processed to give the average prediction quality as a function of size of training set.
- This function can be plotted on a graph, is called the **learning curve** for the algorithm on the particular domain

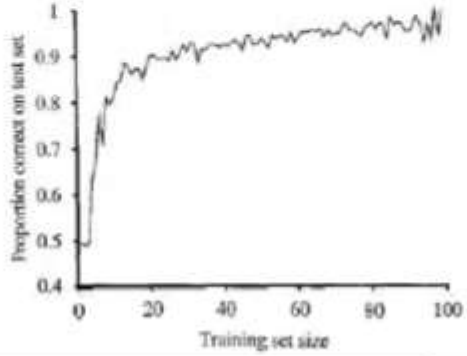


Figure 18.7 A learning curve for the decision tree algorithm on 100 randomly generated examples in the restaurant domain. The graph summarizes 20 trials.

Problem in Decision Trees

Problems in Decision Trees

- Missing data
- Multivalued attributes
- Continuous and integer-valued input attributes
- Continuous-valued output attributes