CSE 473: Machine Learning

Md. Shariful Islam Bhuyan

Recommended Materials

1. Artificial Intelligence: A Modern Approach. Stuart Russell and Peter Norvig [AIMA]
2. Machine Learning. Tom M. Mitchell [ML]
3. Deep Learning. Ian Goodfellow, Yoshua Bengio and Aaron Courville [DL]
4. Provided Blog articles, Videos, Tutorials etc.

Suggestions

1. Class participation is EXTREMELY important
2. There will be 4 (four) class tests
3. NO cheating
4. Brush up Linear Algebra, Probability, Statistics and Calculus (Chapter 2 & 3 [DL])
5. ASK questions
6. Class times

Learning algorithm

A computer program is said to learn from

experience E

with respect to some class of

task T and **performance measure** P

if its

performance at tasks in T

as measured by P

improves with experience E.

Task: Classification

- Specify which of k categories some input belongs to
  - f(article) = {sports, politics, finance}
- Labeled data as experience
  - Era of big data, deep learning
- Accuracy as performance measure
  - Not always appropriate e.g. cancer test

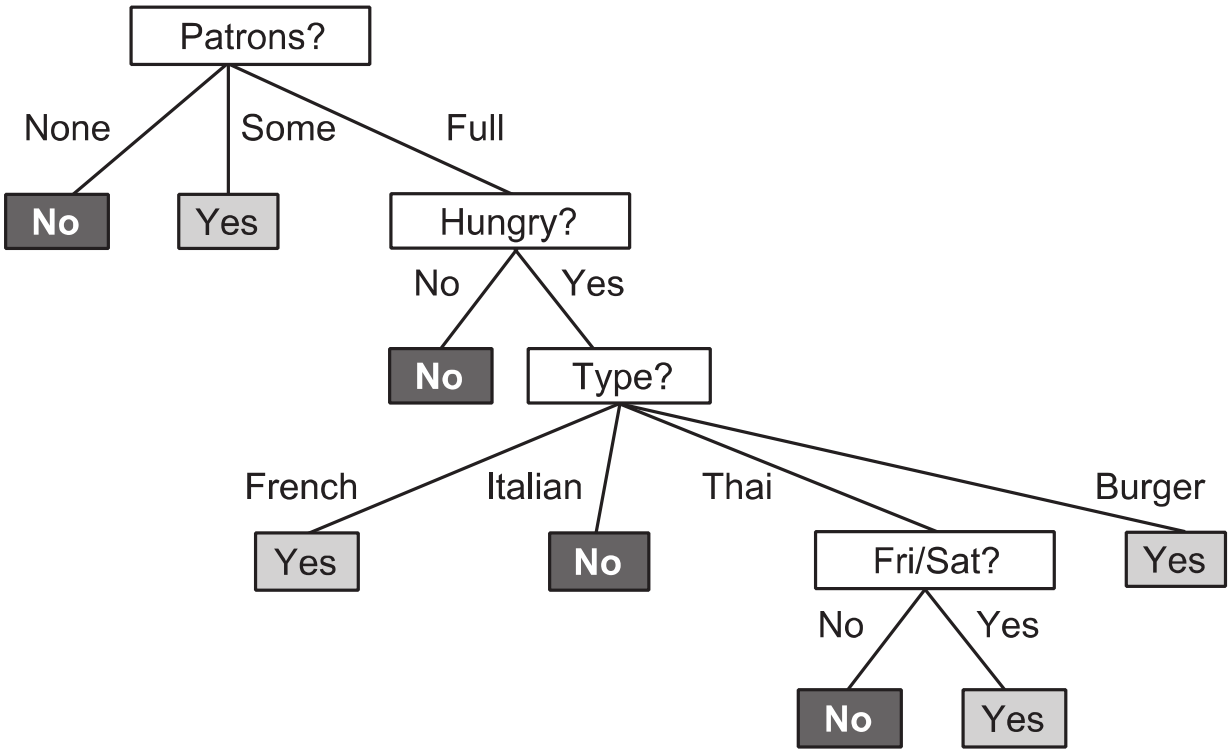
Example: Predict **WillWait**

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, or >60).

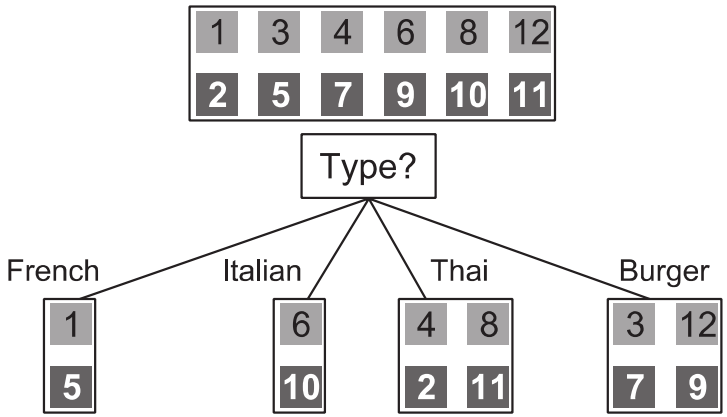
Example: samples

Sample	Alt	Bar	Fri	Hun	Pat	Price	Rain	Res	Type	Est	Goal
x_1	Yes	No	No	Yes	Some	$$$	No	Yes	French	0–10	y_1 = Yes
x_2	Yes	No	No	Yes	Full	$	No	No	Thai	30–60	y_2 = No
x_3	No	Yes	No	No	Some	$	No	No	Burger	0–10	y_3 = Yes
x_4	Yes	No	Yes	Yes	Full	$	Yes	No	Thai	10–30	y_4 = Yes
x_5	Yes	No	Yes	No	Full	$$$	No	Yes	French	>60	y_5 = No
x_6	No	Yes	No	Yes	Some	$$	Yes	Yes	Italian	0–10	y_6 = Yes
x_7	No	Yes	No	No	None	$	Yes	No	Burger	0–10	y_7 = No
x_8	No	No	No	Yes	Some	$$	Yes	Yes	Thai	0–10	y_8 = Yes
x_9	No	Yes	Yes	No	Full	$	Yes	No	Burger	>60	y_9 = No
x_{10}	Yes	Yes	Yes	Yes	Full	$$$	No	Yes	Italian	10–30	y_{10} = No
x_{11}	No	No	No	No	None	$	No	No	Thai	0–10	y_{11} = No
x_{12}	Yes	Yes	Yes	Yes	Full	$	No	No	Burger	30–60	y_{12} = Yes

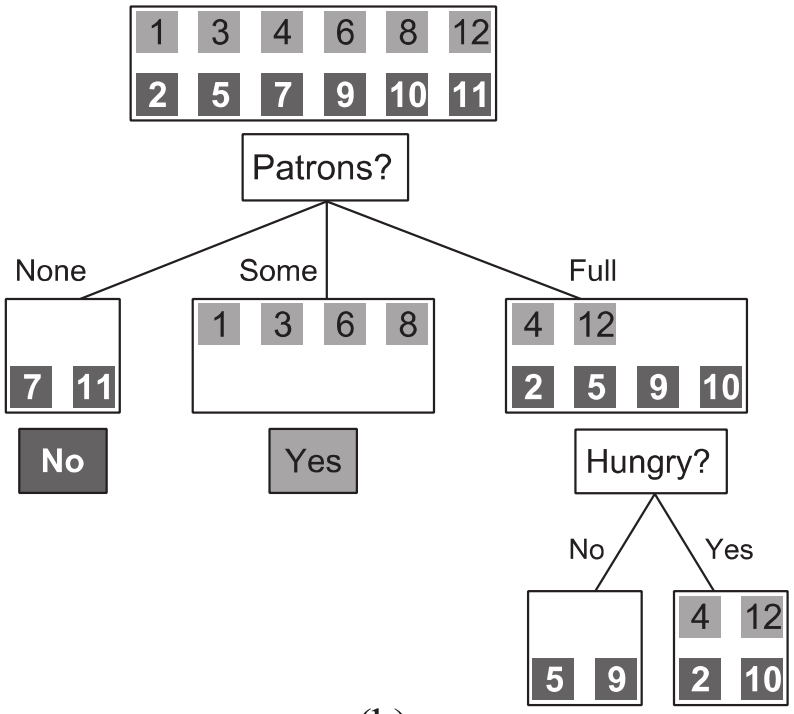
Decision tree



Inducing decision trees from examples



(a)



(b)

Inducing decision trees from examples

- Four cases for subproblems
  1. All positive (or all negative): we are done
  2. Some positive and some negative: choose next **important** attribute (greedy)
  3. No examples left: unobserved case, use prior knowledge/plurality
  4. No attributes left: error, noise, partial information, inherent uncertainty

Decision tree pseudocode

```
function DECISION-TREE-LEARNING(examples, attributes, parent examples) returns a tree
    if examples is empty then return PLURALITY-VALUE(parent examples)
    else if all examples have the same classification then return the classification
    else if attributes is empty then return  PLURALITY-VALUE(examples)
    else
        A ← argmax_{a ∈ attributes} IMPORTANCE(a, examples)
        tree ← a new decision tree with root test A
        for each value v_k of A do
                exs ← { e : e ∈ examples and e.A = v_k }
                subtree ← DECISION-TREE-LEARNING(exs, attributes – A, examples)
                add a branch to tree with label (A = v_k) and subtree subtree
        return tree
```