Applied Machine Learning

Dr. Harikrishnan N B
Computer Science and Information Systems

BITS Pilani
Pilani Campus

SE ZG568 / SS ZG568,
Applied Machine Learning
Lecture No. 7 [02- March-2025]

Decision Tree

Learning - Function Approximation

$x = (x_1, x_2, \cdots x_n) \quad y$

$f: x \rightarrow y$

Problem Setting

- Set of possible instances *X.*
- Unknown target function *f: X—> Y*
- Set of function hypotheses *H* = *{h| h: X—> Y}*

Input

Training Examples of unknown target function *f.*

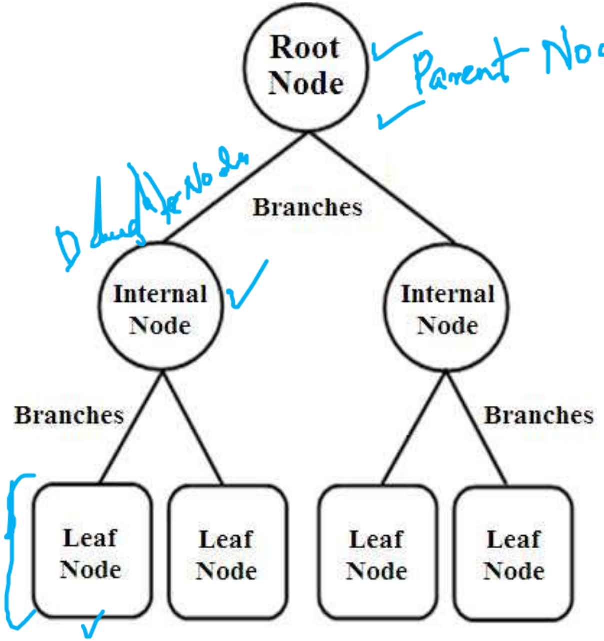
Output

Hypothesis *h* $\in$ *H* that best approximates target function *f.*

Decision Tree

'h' Decision Tree

Decision Tree Learning is a method for approximating the target function **(*Y*)**, in which the **learned functions (h)** is represented by a **decision tree**.



Parent Node

Daughter Node

$h: X \rightarrow Y$

$$\boldsymbol{h: X \text{—}> Y}$$

Dataset $[x_1 \; x_2 \; \ldots \; x_n]$ $[y]$

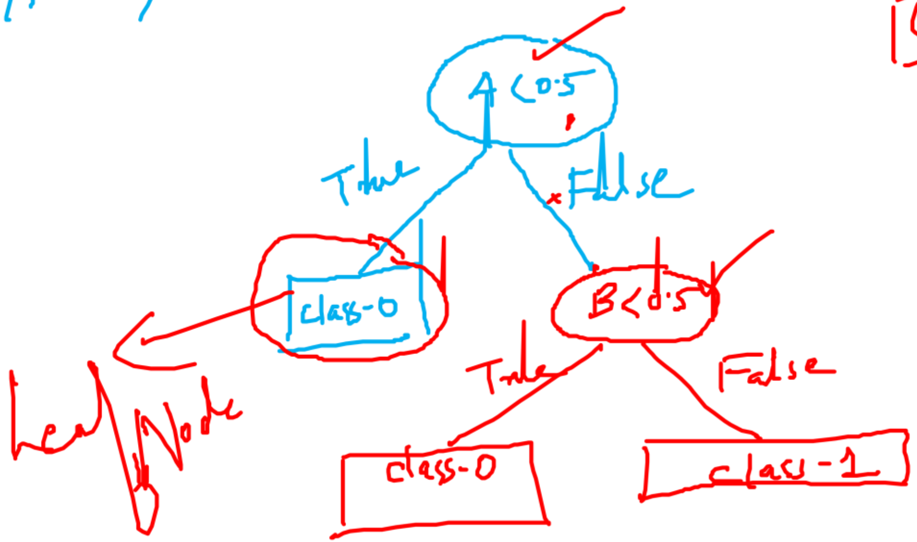
AND

A	B	y	
0	0	0	class-0
0	1	0	
1	0	0	
1	1	1	→ class-1

features A , B

o/p y

class-0
class-1

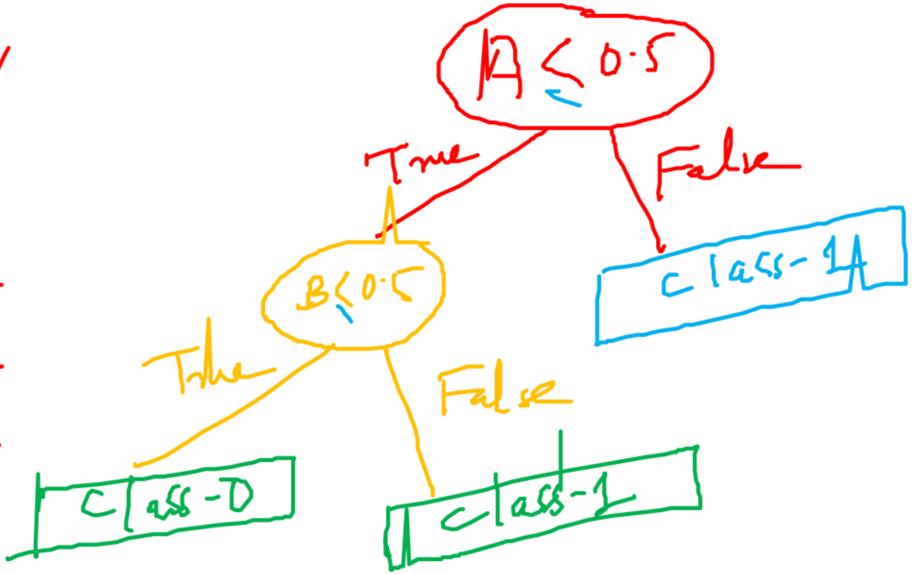


A B
0 0
0 1

if A = 1
  check for B
  if B = 1
    print (o/p as 1)
  else print (o/p as 0)

OR

A	B	Y
0	0	0
0	1	1
1	0	1
1	1	1



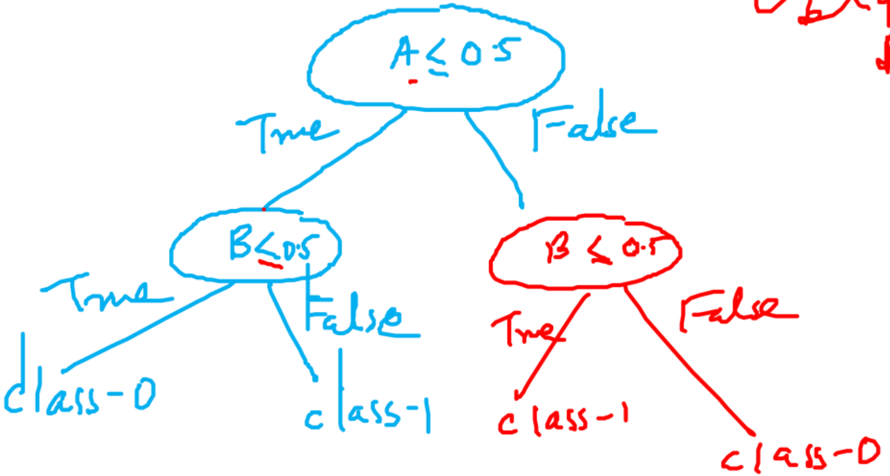
XOR

ID3

feature < threshold

"Oblique Decision Tree"

A	B	Y
0	0	0
0	1	1
1	0	1
1	1	0

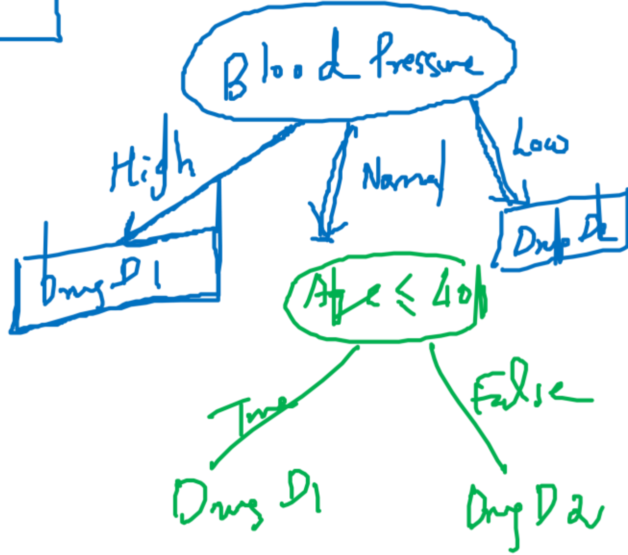


DRUG SUGGESTION EXAMPLE

O/p : Drug

feature: Gender, Age, B.P

Patient ID	Gender	Age	Blood Pressure	Drug
1	M	20	N	D1
2	F	73	N	D2
3	M	37	H	D1
4	M	33	L	D2
5	F	48	H	D1
6	M	29	N	D1
7	F	52	N	D2
8	M	42	L	D2
9	M	61	N	D2
10	F	30	N	D1
11	F	26	L	D2
12	M	54	H	D1



Decision Tree

- **ID3** (Iterative Dichotomiser 3)- **1986 Ross Quinlan**
- **C4.5** - is the successor to ID3 and removed the restriction that features must be categorical by dynamically defining a discrete attribute (based on numerical variables) that partitions the continuous attribute value into a discrete set of intervals.
- **C5.0** - is Quinlan's latest version release under a proprietary license. It uses less memory and builds smaller rulesets than C4.5 while being more accurate.
- **CART** - Classification and Regression Trees - s very similar to C4.5, but it differs in that it supports numerical target variables (regression) and does not compute rule sets. CART constructs binary trees using the feature and threshold that yield the largest information gain at each node.

scikit-learn uses an optimized version of the CART algorithm; however, the scikit-learn implementation does not support categorical variables for now.

https://scikit-learn.org/stable/modules/tree.html#classification

Play Tennis or Not?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

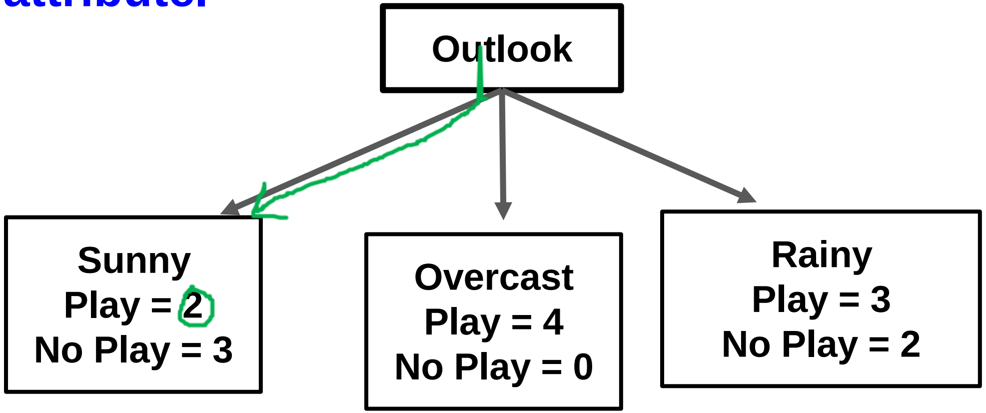
How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

ID3- Decision Tree algorithm for classification

ID3 algorithm

Step 1: Select an attribute to place on the root node and make one branch for each possible value of the attribute.

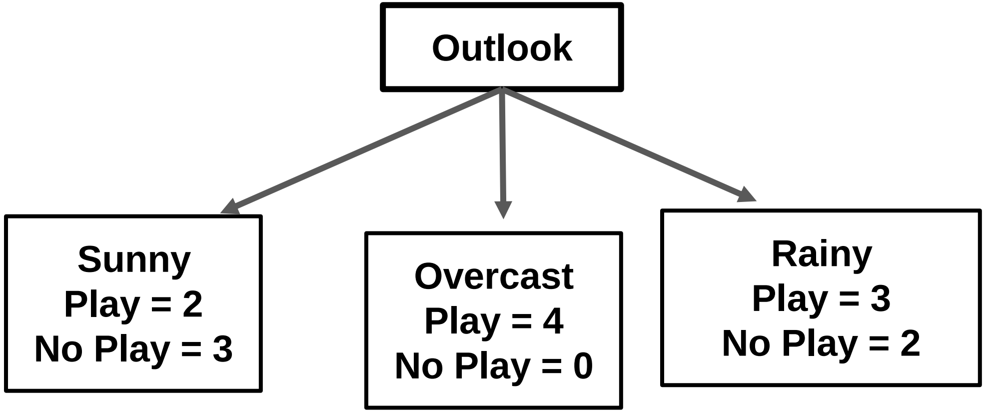


Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

ID3 algorithm

Step 2: Make an assessment of the quality of the split

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

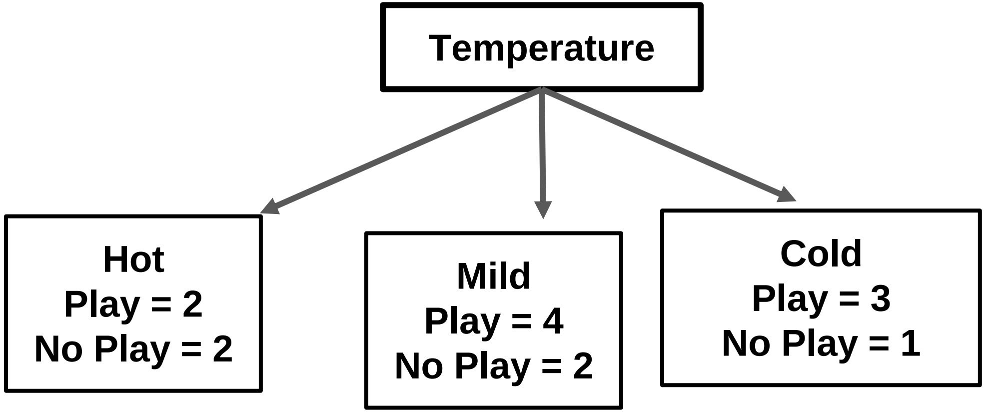


HOW DO YOU MEASURE THE QUALITY?

ID3 algorithm

Step 3: Repeat Step 1 and Step 2 for all other attributes

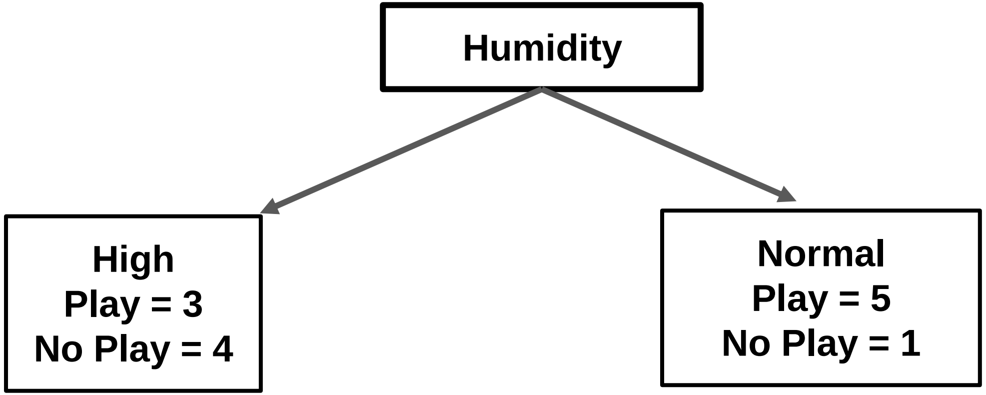
Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No



ID3 algorithm

Step 3: Repeat Step 1 and Step 2 for all other attributes

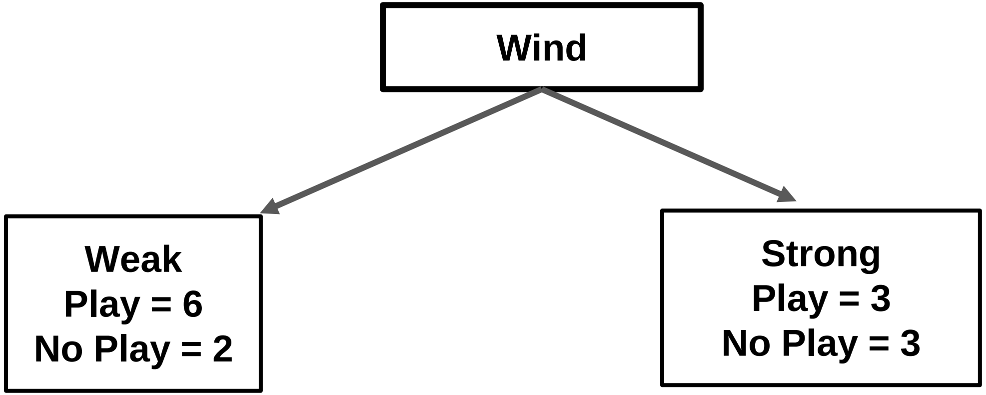
Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No



ID3 algorithm

Step 3: Repeat Step 1 and Step 2 for all other attributes

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

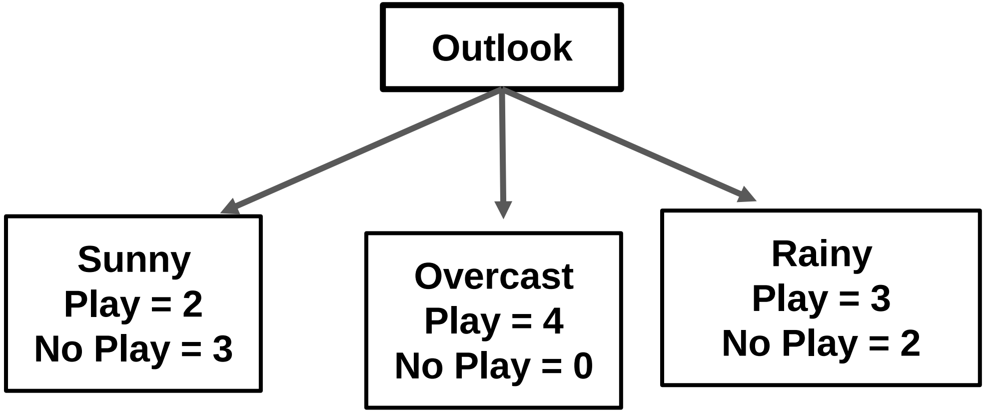


ID3 algorithm

→ Based on the Impurity measure

Step 4: Depending on the QUALITY of the Partial Tree, we select one partial tree

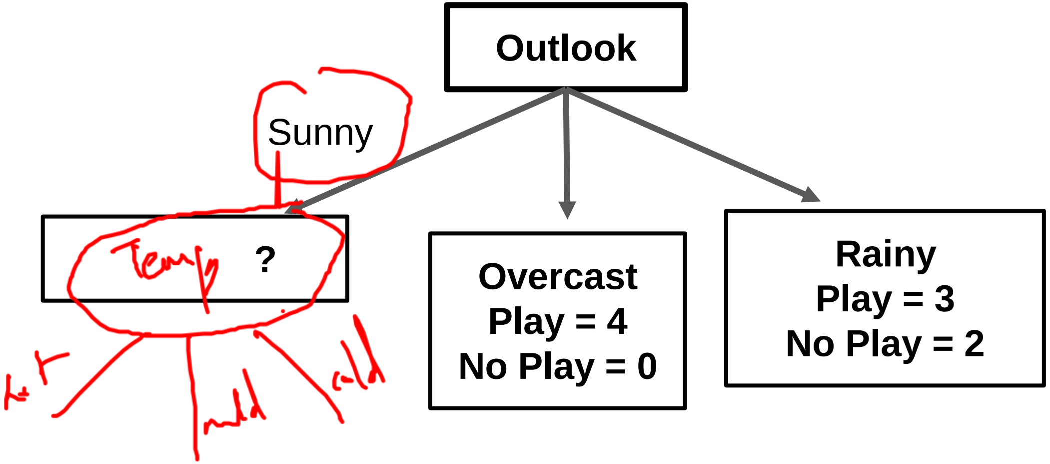
Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No



ID3 algorithm

Step 5: Repeat steps 1 to 4 for each daughter nodes of the selected partial tree

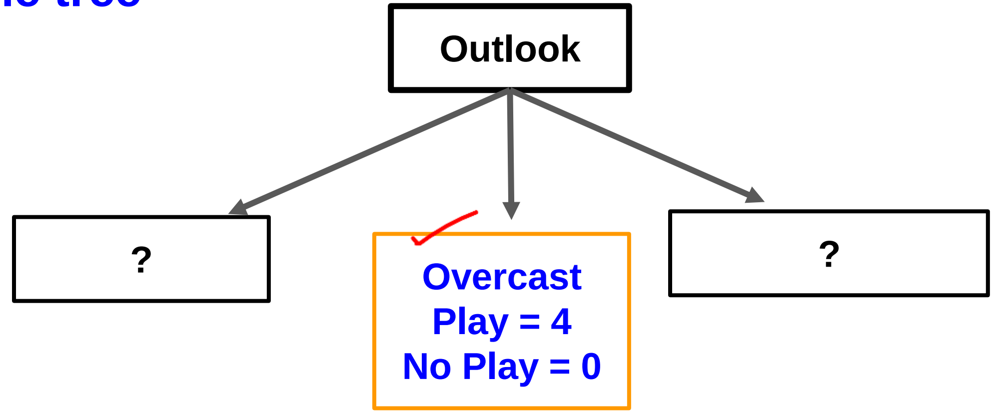
Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No



ID3 algorithm

Step 6: If at any time, all instances at a node have the same classification: stop developing that part of the tree

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

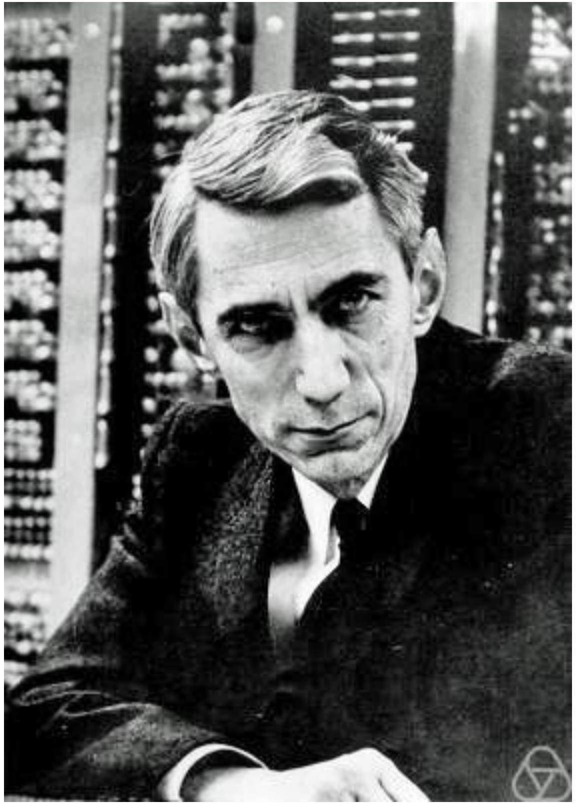


A Mathematical Theory of Comm.

A Basic Introduction to Information Theory

BC	AC
Love	Mass
Beauty	Length
Compassion	Time
(subjective)	(measure)

information



Information as a Measure of Surprise - Intuition

Let's say A is event, "S" is the sample

Self information of this event A $\propto \frac{1}{P(A)}$

$i(P(A)) \propto \frac{1}{P(A)}$

Axiomatic Approach to define self Information

Axiom 1: $i(P(A)) \geq 0$

Axiom 2: If $P(A) \leq P(B)$ then $i(P(A)) \geq i(P(B))$

"i" self Information

Axiom 3: if $A = \phi$

$i(P(A)) \longrightarrow \infty$

$i(P(S)) \Rightarrow 0$

Axiom 4: if A & B are independent

$P(A \cap B) = P(A) \cdot P(B)$

$i(P(A \cap B)) = i(P(A)) + i(P(B))$

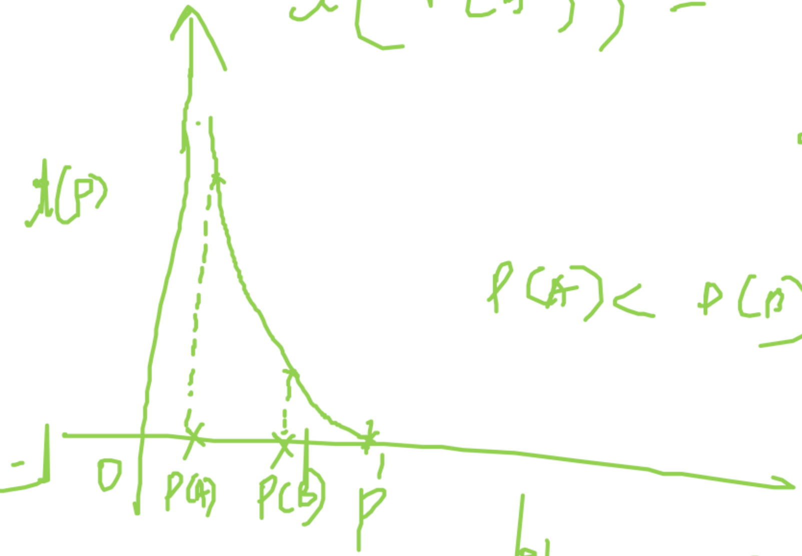
$P(A/B)$

$= \dfrac{P(A \cap B)}{P(B)}$

if A & B are independent $P(A/B) = P(A)$

$$i(P(A)) = -\log(P(A)) \text{ bits}$$

$$i(P(A)) \geq 0$$



$$P(A) < P(B)$$

$$-\log(P(S))$$

$$-\log_2 1 = 0$$

if A & B are independent

$$i(P(A\cap B)) = -\log_2(P(A\cap B))$$

$$= -\log_2(P(A)P(B))$$

$$= -\log_2(P(A)) + -\log_2(P(B))$$

$$i(P(A)) + i(P(B))$$

What is the average surprise or average self information for the Random Experiment.

$$E(x) = \sum_{i=1}^{n} x_i \, p(x_i)$$

$$E\left[i\left(p(x) \right) \right]$$

$$\sum_{x \in S} p(x) \, i\left(p(x) \right)$$

$$= \sum p(x) \left(-\log_2 \left(p(x) \right) \right)$$

$$\boxed{H(x) = -\sum_{x \in S} p(x) \log_2 p(x)}$$

entropy

unit : bits

Coin Toss experiment

$S = \{H, T\}$ $\quad P(H) = 1/2, P(T) = 1/2$

$i(P(H)) = -\log_2\left(\frac{1}{2}\right) = 1\text{ bit}$ $\qquad i(P(T)) = -\log_2\left(\frac{1}{2}\right) = 1\text{ bit}$

$E[X] = \sum_{i=1}^{N} x_i\, p(x_i)$

H

$E[i(P(X))] = \sum_{j=1}^{n} \underbrace{i(P(x_j))}\, P(x_j)$

$-\sum_{j=1}^{n} \log_2(P(x_j)) \times P(x_j)$

$-P(H) \times \log_2(P(H)) - P(T)\log_2(P(T))$

$= +1/2 \times 1 + 1/2 \times 1$

$= 1\text{ bit}$

$\{x_1, x_2\}$

$\downarrow \quad \downarrow$

$P_1 \quad P_2$

$P_1 + P_2 = 1 \Rightarrow P_2 = 1 - P_1$

$P_1 = 1$

N Job offers $J_1 \; J_2 \; \ldots \; J_n$

$P_1 \; P_2 \; \ldots \; P_n$

$P_1 = P_2 = \ldots P_n = \frac{1}{n}$

$-P_1 \log P_1 - P_2 \log P_2 - \ldots - P_n \log P_n$

$- \frac{1}{N} \sum_{i=1} \log P_i$

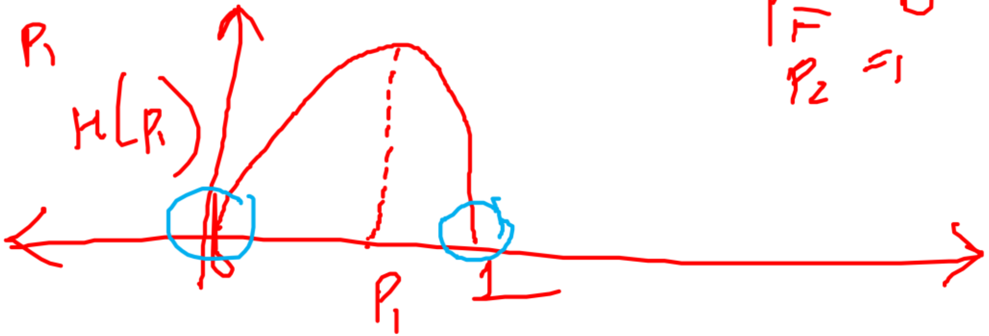
$$H(P_1) = -P_1 \log_2 P_1 - (1 - P_1) \log_2 (1 - P_1)$$

$1 * \log_2 1 \qquad 0 * \log_2 0$

$P_F = 0 \qquad P_1 = 1$

$P_2 = 1 \qquad P_2 = 0$

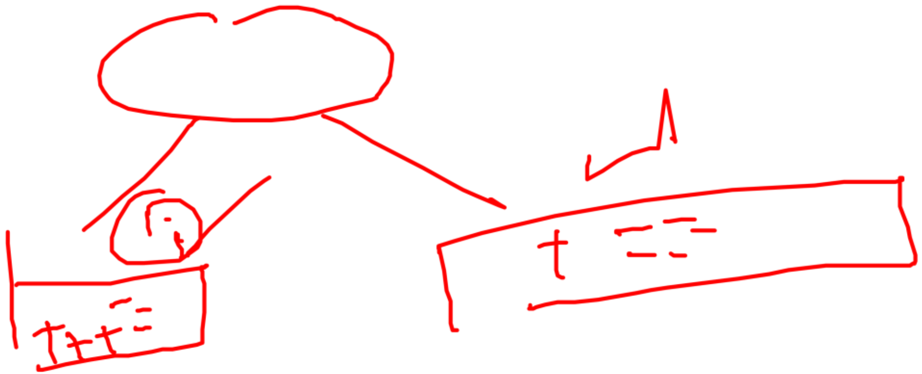
$H(P_1)$ vs P_1



$$-\sum_{i=1}^{n} p_i \log p_i$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{1}{n}\right) = -\frac{1}{\cancel{n}} \times \cancel{n}\log\left(\frac{1}{n}\right)$$

$$= -\log\left(\frac{1}{n}\right) = \underline{\underline{\log_2 n \text{ bits}}}$$



How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

How do we construct such a tree?

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

Decision Tree