
INSTITUT TEKNOLOGI BANDUNG
1920

Modul : Decision Tree Learning (DTL)

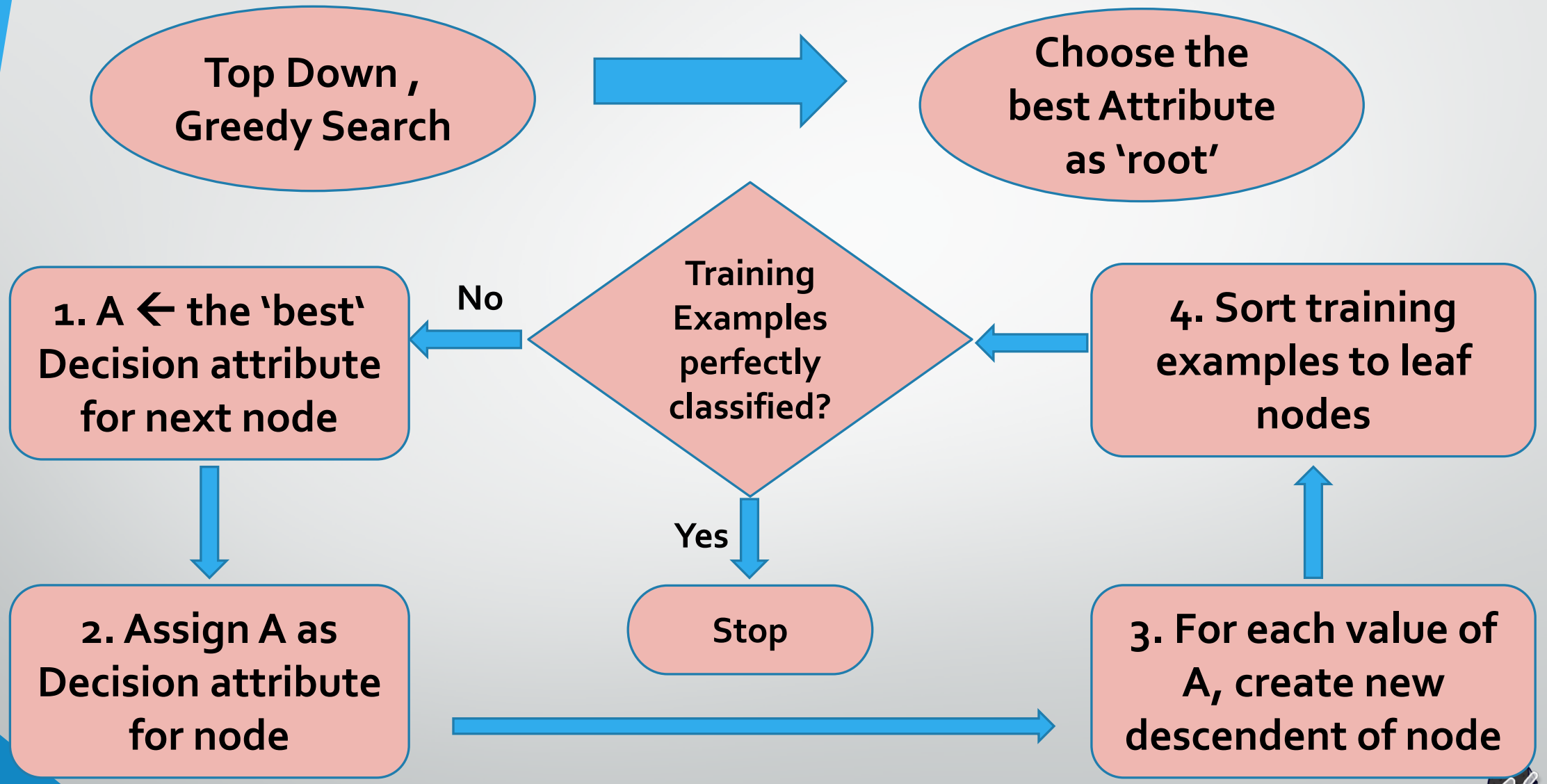
Basic DTL Algorithm (ID3)

Nur ULFA Maulidevi

KK IF - Teknik Informatika- STEI ITB

Pembelajaran Mesin
(Machine Learning)

ID3 Algorithm

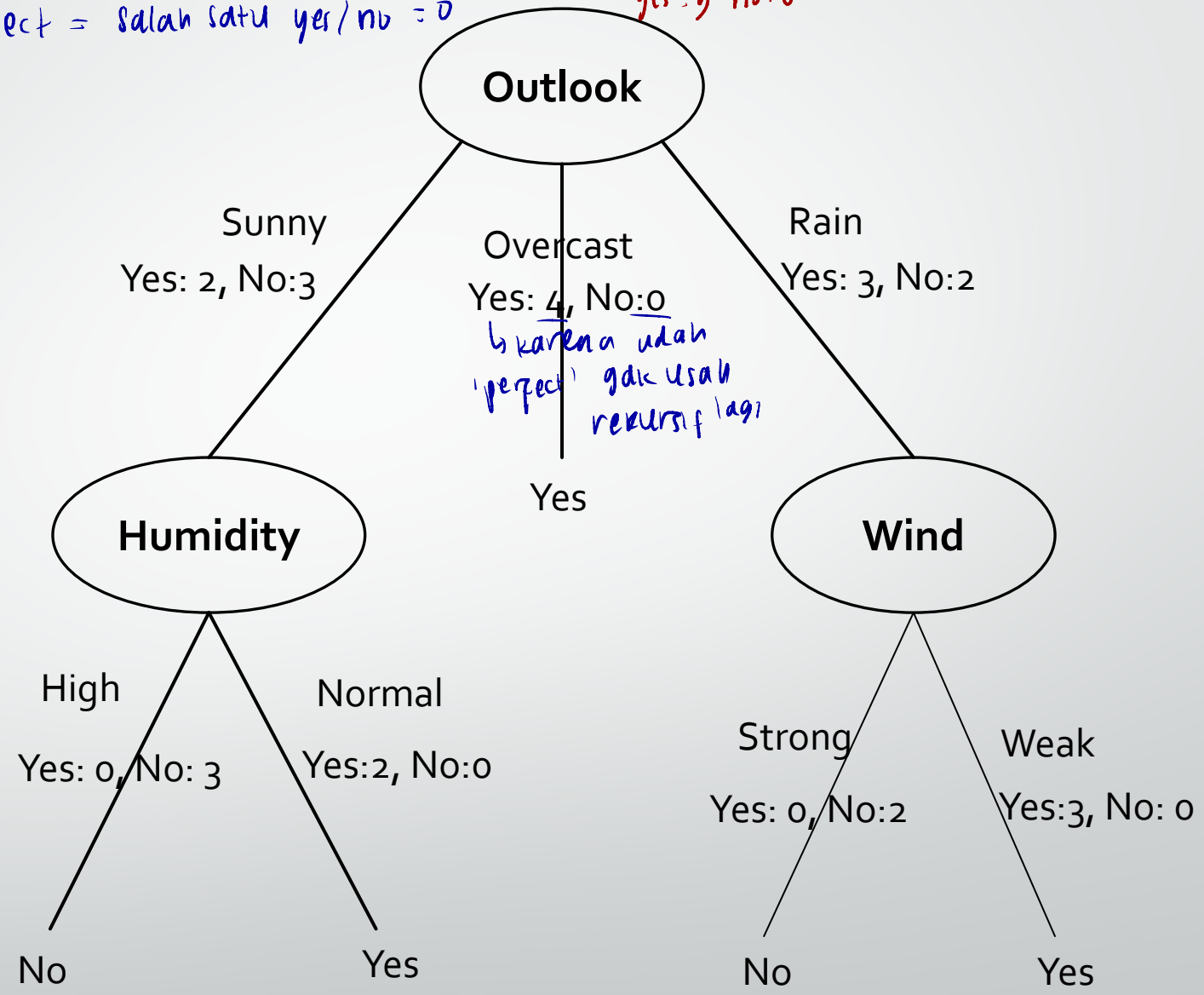


Example

perfect = salah satu yes/no = 0

yes : 9 no : 5

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No



Contoh DTL

DTL (examples, attributes, parent examples)

DTL ($S_{y=9, n=5}$, {O, t, h, w}, -

A ← outlook (cari plg tinggi)

for each $v_k \in$ {sunny, outcas, rain}

$S_{sunny} \leftarrow \{e : e.outlook = sunny\}$

Subtree ← DTL ($S_{y=2, n=3}$, {t, h, w}, $S_{y=9, n=5}$)

A ← humidity → dari information gain

for each $v_t \in$ {high, normal}

$S_{high} \subset \{e : e.humidity = high\}$

Subtree ← DTL {$S_{y=0, n=2}$, {t, w}, $S_{y=2, n=3}$}

≠ leaf N

Temperatur : importancenya kurang

Decision Tree Learning (Russel & Norvig, 2021)

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
  $A \leftarrow \text{argmax}_{a \in attributes}$ IMPORTANCE(a, *examples*)
  tree ← a new decision tree with root test A
  for each value v_k of A **do**
    $exs \leftarrow \{e : e \in examples \text{ and } e.A = v_k\}$
    subtree ← DECISION-TREE-LEARNING(exs, *attributes* − A, *examples*)
    add a branch to *tree* with label $(A = v_k)$ and subtree *subtree*
  return *tree*

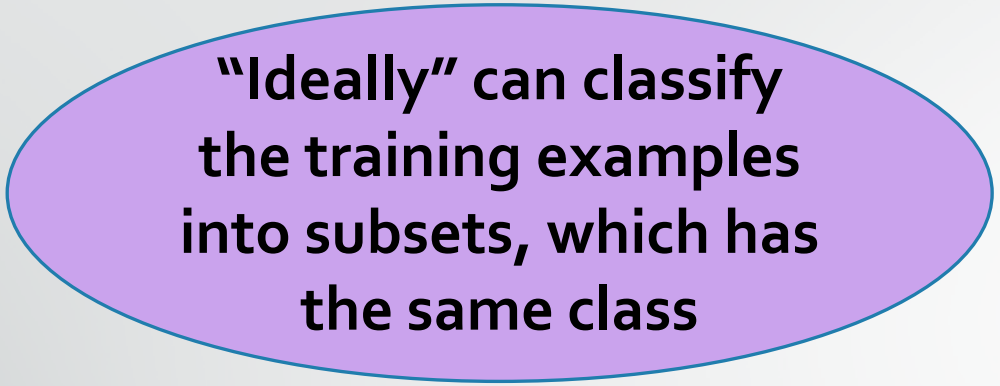
Information gain
Gain ratio

|A|=2: binary tree
|A| min 2: n-ary tree

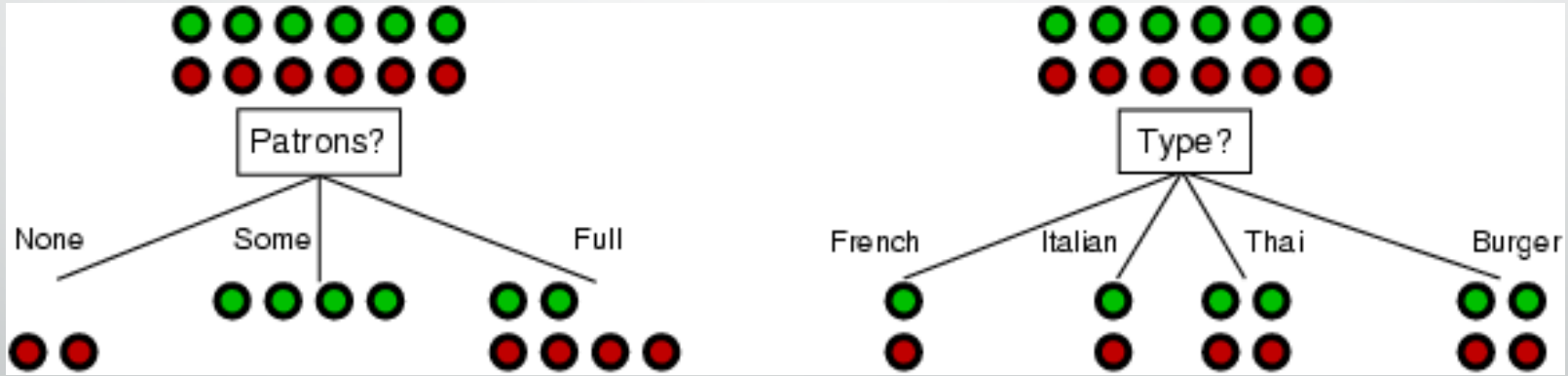
The function PLURALITY-VALUE selects the most common output value among a set of examples, breaking ties randomly.

Best Attribute

"Ideally" can classify the training examples into subsets, which has the same class



Information Gain



THANK YOU


INSTITUT TEKNOLOGI BANDUNG
1920