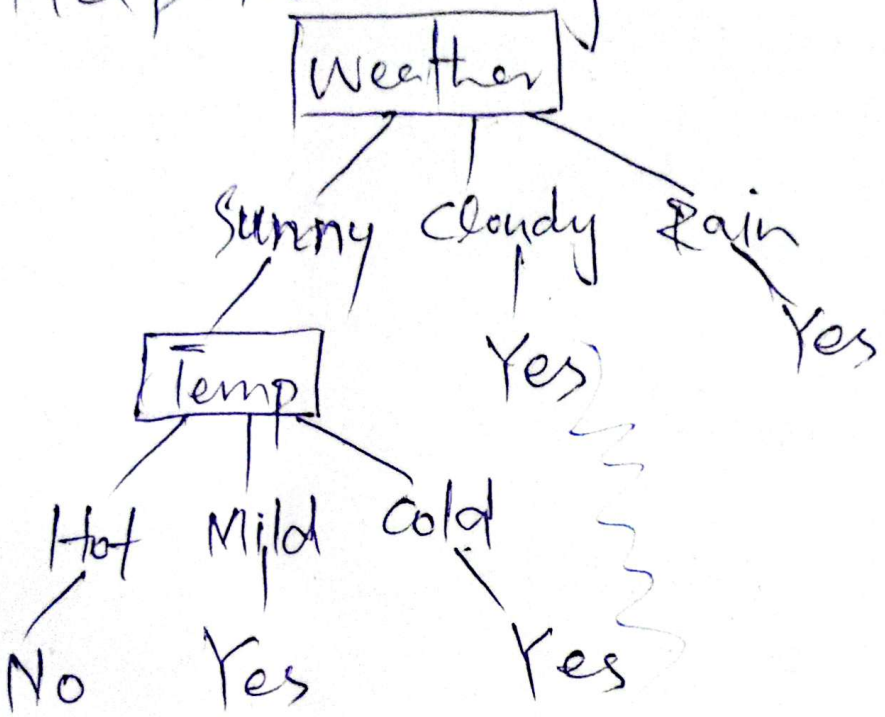
~~Dummy Variable~~

Decision Tree

→ Logrithm used for both classification and regression tasks

→ Help in taking decision



→ Tree Structure
→ Decision node
→ leaf node
→ Splitting
→ Entropy and information gain
→ Pruning

Drawing decision Tree :-

Day	Weather	Temp	Humi	Wind	Foot
Day1	Sunny	Hot	High	Weak	No
Day2	Sunny	Hot	〃	Strong	No
Day3	Cloudy	Hot	〃	Weak	Yes
Day4	Rain	Mild	〃	Weak	Yes
Day5	Sunny	Cool	Normal	Weak	Yes
Day6	〃	Mild	〃	Weak	No
Day7	〃	Mild	High	Strong	Yes
Day8	Cloudy	Mild	Normal	Weak	No

Calculation :-

→ Calculate IG of weather

→ Step 1 :- Entropy of entire dataset.

$$S\{+4,-4\} = -\frac{4}{8}\log_2\frac{4}{8} - \frac{4}{8}\log_2\frac{4}{8}$$

→ Step 2 :- Entropy of all attributes

Entropy of Sunny $\{+2,-3\}$
〃 〃 Cloudy $\{+1,-1\}$
〃 〃 Rain $\{+1,0\}$

Information Gain = Entropy(W.D) − S(Sunny) − S(Cloudy) − S(Rain)

- Calculate IG of temperature
- Calculate IG of wind

- Gain(S, Weather) = 0.246 → Highest
- Gain(S, Temperature) = 0.029
- Gain(S, Humidity) = 0.15
- Gain(S, wind) = 0.0478