10/8/2025 Lab 2

* Implement ID3 algorithm using decision tree method
Use appropriate dataset for building a decision tree and apply this knowledge to classify the example

Dataset

Day	Outlook	Temp	Humidity	Wind	PlayTennis
D1	Sunny-n	Hot	high	weak	No
D2	Sunny-n	Hot.	high	Strong	No
D3	Overcast	Hot.	high	weak	Yes
D4	Rain-y	Mild.	high	weak	Yes
D5	Rain-y	Cool	Normal	weak	Yes
D6	Rain-n	Cool	Normal	strong	No
D7	Overcast	Cool	Normal	strong	Yes
D8	Sunny-n	Mild.	High	weak	No
D9	Sunny-y	Cool	normal	weak	Yes
D10	Rain	Mild	normal	weak	Yes
D11	Sunny-y	Mild	Normal	strong	Yes
D12	Overcast	Mild	High	strong	Yes
D13	overcast	Hot.	Normal	weak	Yes
D14	Rain	Mild	High	strong	No

ans Calculate information gained
root node = highest info

Out of 4 attribute, attribute with max = root.

Algorithm

1) Calculate Entropy
Entropy measures the uncertainty in a dataset.
Formula $H(S) = -\sum P_i \log_2 P_i$
where P_i is the probability of class i the dataset S.

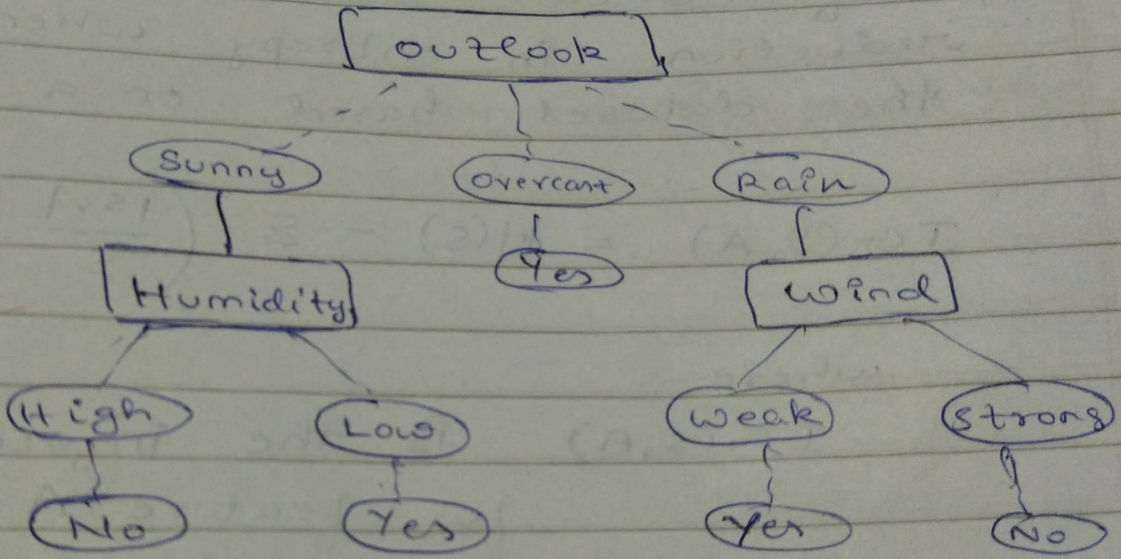
2) Compute information Gain for each feature
Information Gain measures the reduction in entropy when splitting the dataset based on a feature.

$$IG(S, A) = H(S) - \sum \left(\frac{|S_v|}{S} H(S_v) \right)$$

where
$IG(S,A)$ is the information gain for feature A
S_v is the subset where feature A has value 'v'
$H(S)$ is entropy before splitting
$H(S_v)$ is entropy after splitting

3) Select the Best Feature:
Choose the feature with the highest information gain as the root node.

Final Decision Tree



4) Split the Dataset
Divide the dataset based on the selected feature

5) Recursirely Build the Tree
Repeat steps 1-4 on each subset until
- All instances in a subset belong to same class
- There are no more features left to split.

6) Base case
- If all instances Belong to the same class, return that class as a leaf node.
- If no features are left, return the majorit class in the subset.

10/3/25