**Decision tree based ID3 algorithm**

To understand the implementation, let us try to implement it to a smaller data set with a bunch of examples to decide if a person wants to go for a walk.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 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 |

Following terminologies are used in this algorithm

 Entropy : Entropy is a measure of impurity

It is defined for a binary class with values a/b as:

 Information Gain : measuring the expected reduction in Entropy

THE PROCEDURE

1) In the ID3 algorithm, begin with the original set of attributes as the root node.

2) On each iteration of the algorithm, iterate through every unused attribute of the remaining set and calculates the entropy (or information gain) of that attribute.

3) Then, select the attribute which has the smallest entropy (or largest information gain) value. 4) The set of remaining attributes is then split by the selected attribute to produce subsets of the data.

5) The algorithm continues to recurs on each subset, considering only attributes never selected before.

**Dataset Details**

playtennis dataset which has following structure

Total number of instances=15

Attributes=Outlook, Temperature, Humidity, Wind, Answer

Target Concept=Answer

