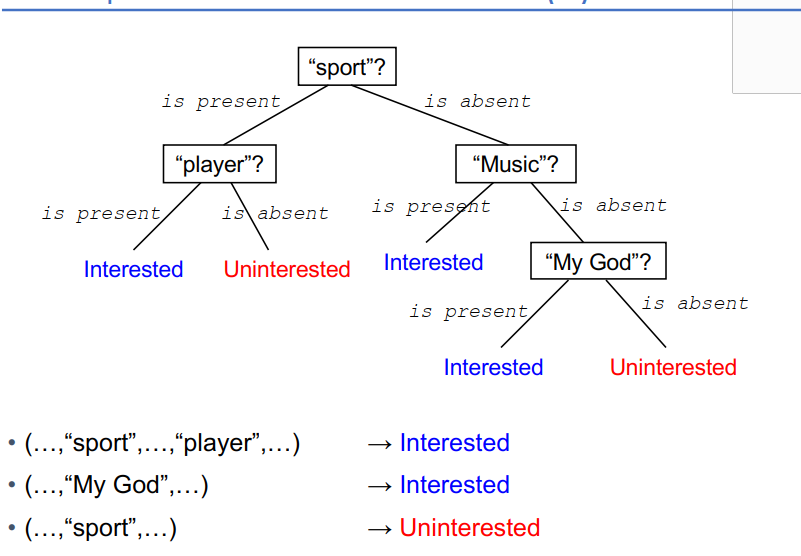
Decision tree: Dùng để biểu diễn 1 func bằng cách sử dụng Tree

+Mỗi cây quyết định có thể hiểu là 1 tập của các IF-Then

* Cây quyết định được sử dụng nhiều trong các ứng dụng thực tế

***Ex:*** ******

Diagram

Description automatically generated

Mỗi nút bên trong đại diện cho một thuộc tính để kiểm tra dữ liệu đến.

Mỗi nhánh/cây con của một nút tương ứng với một giá trị của thuộc tính của nút đó.

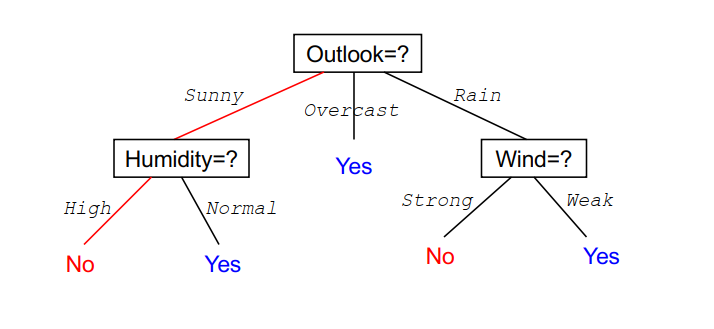
Mỗi nút lá đại diện cho một nhãn lớp.

Khi một cây đã được học, chúng ta có thể dự đoán nhãn cho một phiên bản mới bằng cách sử dụng các thuộc tính của nó để di chuyển từ gốc xuống một chiếc lá.

¨ Nhãn của lá sẽ được sử dụng để gán cho thể hiện mới.

Mỗi đường dẫn từ gốc đến lá là sự kết hợp/VÀ của các bài kiểm tra thuộc tính.

Bản thân cây quyết định là sự phân tách/HOẶC của những liên từ.



Diagram

Description automatically generated

**Học = ID3**

ID3 (Iterative Dichotomiser 3) là một thuật toán tham lam mà được đề xuất bởi Ross Quinlan vào năm 1986.

Nó sử dụng sơ đồ từ trên xuống.

Tại mỗi nút N, chọn thuộc tính kiểm tra A có thể giúp chúng tôi tốt nhất là phân loại cho dữ liệu trong N.

¨ Tạo một nhánh cho mỗi giá trị của A, sau đó tách các dữ liệu vào các nhánh của nó cho phù hợp.

Phát triển cây cho đến khi:

¨ Nó phân loại chính xác tất cả dữ liệu huấn luyện; hoặc là

¨ Tất cả các thuộc tính được sử dụng.

*Lưu ý:* mỗi thuộc tính chỉ được xuất hiện nhiều nhất một lần trong bất kỳ

đường đi của cây.

**ID3\_alg**(*Training\_Set, Class\_Labels, Attributes*)  
Generate the Root of the tree  
If all of *Training\_Set* belong to class c, then Return Root as leaf with label c  
If *Attributes* is empty, then  
Return Root as leaf with label c = **Majority\_Class\_Label**(*Training\_Set*)  
A ← a set of *Attributes* that are best discriminative for *Training\_Set*Let A be the test attributes of Root  
For each value v of A  
Generate a branch of Root which corresponds with v.  
Determine Training\_Setv = { x in *Training\_Set* | xA = v}  
If (Training\_Setv is empty) Then  
Generate a leaf with class label c = **Majority\_Class\_Label**(*Training\_Set*)  
Else  
Generate a subtree by **ID3\_alg**(Training\_Setv, Class\_Labels, Attributes \{A})  
Return Root