Árbol de clasificación

Dado un conjunto D de datos etiquetados el problema es construir un árbol que permita mediante preguntas sobre los atributos permita asignar clase a un patrón, dato no etiquetado. La idea es conseguir lo anterior con el mínimo número de preguntas, con la mayor generalidad posible.

Para lograr lo anterior en cada etapa de la construcción del árbol se elige para la pregunte al atributo que máximiza la ganancia de información dada por

$G(H(D), a) = H(D) - \sum_i P_i H(D_i)$, $D = \bigcup D_i$, D_i forman una partición de D preguntando por los valores del atributo a,

$P_i = \frac{|D_i|}{|D|}$ y $H(D) = -\sum_{k \in C} P_k \log(P_k)$, k indice de clase.

$H(D)$ es la entropia de los datos de acuerdo a la clase a que pertenecen, cuando todos son de la misma clase la entropia es cero es decir estan totalmente en orden de acuerdo a las clases. De acuerdo a la anterior Euristica el algoritmo de entrenamiento funciona como sigue:

```
create_Branch (Datos):
    checar si cada dato es de la misma clase:
        Si es el caso regresa la etiqueta de clase
        Else
            Encontrar el mejor atributo para partir los datos
            Partir los datos mediante la pregunta
            Crear nodo correspondiente
            Para cada partición Di
                create_Branch (Di) y conectar el resultado
            regresa el nodo creado
```