**Áp dụng thuật toán ID3 để xây dựng cây cho dữ liệu sau  
Cách 1: Tính Entropy và Information Gain (IG) để xây dụng cây**

* Entropy(S) =
* **Age** = {<=30 ; 30…40 ; >40}

~ {2+/3- ; 4+/0- ; 3+/2-}

Entropy () = - 2/5log2(2/5) – 3/5log2(3/5)

- 3/5log2(3/5) – 2/5log2(2/5)

- 4/4log2(4/4) – 0/4log2(0/4)

* **Income** = {low; medium; high}

~ {3+/1- ; 4+/2-; 2+/2-}

Entropy () = - 2/4log2(2/4) – 2/4log2(2/4)

- 4/6log2(4/6) – 2/6log2(2/6)

- 3/4 log2(3/4) – 1/4 log2(1/4)

* **Student** = {No ; Yes}

~ {3+/4-; 6+/1-}

Entropy () = - 3/7 log2(3/7) – 4/7 log2(4/7)

- 6/7 log2(6/7) – 1/7 log2(1/7)

* **Credit Rating** = {Fair ; Excellent}

~ {6+/2-; 3+/3-}

Entropy () = - 3/6 log2(3/6) – 3/6log2(3/6)

- 6/8log2(6/8) – 2/8log2(2/8)

**Cách 2: Dùng phần mềm Weka, Orange hoặc Visual Studio.**

**Phần mềm Orange**



