

تمرین

■ درخت تصمیم را برای داده های زیر رسم کنید :

Table 1.1 • Hypothetical Training Data for Disease Diagnosis

Patient ID#	Sore Throat	Fever	Swollen Glands	Congestion	Headache	Diagnosis
1	Yes	Yes	Yes	Yes	Yes	Strep throat
2	No	No	No	Yes	Yes	Allergy
3	Yes	Yes	No	Yes	No	Cold
4	Yes	No	Yes	No	No	Strep throat
5	No	Yes	No	Yes	No	Cold
6	No	No	No	Yes	No	Allergy
7	No	No	Yes	No	No	Strep throat
8	Yes	No	No	Yes	Yes	Allergy
9	No	Yes	No	Yes	Yes	Cold
10	Yes	Yes	No	Yes	Yes	Cold

Step 1:

$$\text{Ent}(\text{Class}) = (-3/10)\log(3/10) + (-3/10)\log(3/10) + (-4/10)\log(4/10) = 0.472$$

$$\text{Ent}(\text{ST}+) = (-2/5)\log(2/5) + (-1/5)\log(1/5) + (-2/5)\log(2/5) = 0.458$$

$$\text{Ent}(\text{ST}-) = (-2/5)\log(2/5) + (-1/5)\log(1/5) + (-2/5)\log(2/5) = 0.458$$

$$\text{Gain}(\text{ST}) = 0.472 - 5/10(0.458) - 5/10(0.458) = 0.014$$

$$\text{Ent}(\text{F}+) = (-1/5)\log(1/5) + (-4/5)\log(4/5) = 0.217$$

$$\text{Ent}(\text{F}-) = (-2/5)\log(2/5) + (-3/5)\log(3/5) = 0.292$$

$$\text{Gain}(\text{ST}) = 0.472 - 5/10(0.217) - 5/10(0.292) = 0.218$$

$$\text{Ent}(\text{SG}+) = (-3/3)\log(3/3) = 0$$

$$\text{Ent}(\text{SG}-) = (-4/7)\log(4/7) + (-3/7)\log(3/7) = 0.296$$

$$\text{Gain}(\text{SG}) = 0.472 - 3/10(0) - 4/10(0.296) = 0.353$$

$$\text{Ent}(\text{C}+) = (-1/8)\log(1/8) + (-3/8)\log(3/8) + (-4/8)\log(4/8) = 0.423$$

$$\text{Ent}(\text{C}-) = 0$$

$$\text{Gain}(\text{C}) = 0.472 - 8/10(0.423) = 0.133$$

$$\text{Ent}(\text{H}+) = (-1/5)\log(1/5) + (-2/5)\log(2/5) + (-2/5)\log(2/5) = 0.458$$

$$\text{Ent}(\text{H}-) = (-1/5)\log(1/5) + (-2/5)\log(2/5) + (-2/5)\log(2/5) = 0.458$$

$$\text{Gain}(\text{H}) = 0.472 - 0.458 = 0.014$$

Step 2:

$$\text{Ent}(\text{SG}-) = 0.296$$

$$\text{Ent}(\text{ST}|\text{SG}-) = (-2/4)\log(2/4) + (-2/4)\log(2/4) = 0.301$$

$$\text{Ent}(\text{ST}+|\text{SG}-) = (-1/3)\log(1/3) + (-2/3)\log(2/3) = 0.276$$

$$\text{Gain}(\text{ST}|\text{SG}-) = 0.296 - 4/7(0.301) - 3/7(0.276) = 0.006$$

$$\text{Ent}(\text{F}|\text{SG}-) = 0$$

$$\text{Ent}(\text{F}+|\text{SG}-) = 0$$

$$\text{Gain}(\text{F}|\text{SG}-) = 0.296 - 0 = 0.296$$

$$\text{Ent}(\text{C}|\text{SG}-) = 0$$

$$\text{Ent}(\text{C}+|\text{SG}-) = (-4/7)\log(4/7) + (-3/7)\log(3/7) = 0.296$$

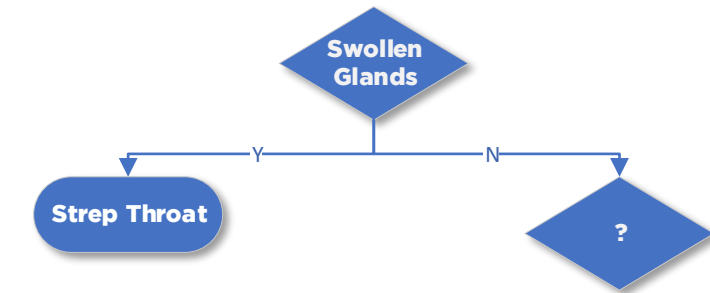
$$\text{Gain}(\text{C}|\text{SG}-) = 0.296 - 0.296 = 0$$

$$\text{Ent}(\text{H}|\text{SG}-) = (-2/4)\log(2/4) + (-2/4)\log(2/4) = 0.301$$

$$\text{Ent}(\text{H}+|\text{SG}-) = (-1/3)\log(1/3) + (-2/3)\log(2/3) = 0.276$$

$$\text{Gain}(\text{H}|\text{SG}-) = 0.296 - 4/7(0.301) - 3/7(0.276) = 0.006$$

Step 1:



Step 2:

