```
/Users/haim/TEACH/ML/online/Q/id3-02.stex
% mte05 ho07 h008
```

You are given the following training data:

	Height	Hair	Eyes	Sensitivity
1	short	blond	blue	yes
2	tall	blond	brown	no
3	tall	red	blue	yes
4	tall	dark	brown	no
5	short	dark	blue	no
6	tall	dark	blue	no
7	tall	blond	blue	yes
8	short	blond	brown	no

The target attribute is Sensitivity.

Compute the decision tree using the ID3 algorithm.

Answer

$$
\begin{aligned}
s_0 &= \{1,2,3,4,5,6,7,8\} \\
E(s_0) &= 0.95443 \\
E(s_0|\text{Height}) &= 0.9512, & Gain(s_0,\text{Height}) &= 0.00323 \\
E(s_0|\text{Hair}) &= 0.5, & Gain(s_0,\text{Hair}) &= 0.4544 \\
E(s_0|\text{Eyes}) &= 0.6068, & Gain(s_0,\text{Eyes}) &= 0.3476
\end{aligned}
$$

So best choice is Hair.
Subtrees "red", "dark" have 0 entropy. We need to split the subtree generated by "blond".

$$
\begin{aligned}
s_1 &= \{1,2,7,8\} \\
E(s_1) &= 1 \\
E(s_1|\text{Height}) &= 1, & Gain(s_1,\text{Height}) &= 0 \\
E(s_1|\text{Eyes}) &= 0, & Gain(s_1,\text{Eyes}) &= 1
\end{aligned}
$$

So best choice is Eyes.