Assume that you have the following training examples available:

<table>
<thead>
<tr>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>p</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>t</td>
<td>t</td>
<td>f</td>
<td>p</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>f</td>
<td>t</td>
<td>f</td>
<td>p</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>t</td>
<td>f</td>
<td>t</td>
<td>p</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>n</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>f</td>
<td>t</td>
<td>t</td>
<td>n</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>n</td>
</tr>
</tbody>
</table>

Use all of the training examples to construct a decision tree. In case of ties between features, break ties in favor of features with smaller numbers (for example, favor F1 over F2, F2 over F3, and so on).

How does the resulting decision tree classify the following example:

<table>
<thead>
<tr>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>f</td>
<td>f</td>
<td>t</td>
<td>t</td>
<td>?</td>
</tr>
</tbody>
</table>
Some formula:

\[
\text{Gain}(A) = I(p,n) - E(A)
\]

\[
E(A) = \frac{p_t + n_t}{p+n} I(p_t + n_t) + \frac{p_f + n_f}{p+n} I(p_f + n_f)
\]

\[
I(p,n) = -\frac{p}{p+n} \log_2 \frac{p}{p+n} - \frac{n}{p+n} \log_2 \frac{n}{p+n}
\]

Pre-compute some \(I(x,y)\)

\[
I(0,x) = I(x,0) = -1 \log_2 1 - 0 \log_2 0 = 0
\]

\[
I(x,x) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 1
\]

\[
I(1,2) = I(2,1) = -\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3} = 0.918
\]

\[
I(1,3) = I(3,1) = -\frac{3}{4} \log_2 \frac{3}{4} - \frac{1}{4} \log_2 \frac{1}{4} = 0.811
\]

\[
I(3,4) = I(4,3) = -\frac{4}{7} \log_2 \frac{4}{7} - \frac{3}{7} \log_2 \frac{3}{7} = 0.985
\]

First, choose from \{F1, F2, F3, F4, F5\} to become the root.

\[
\begin{array}{ccccccc}
F1 & F2 & F3 & F4 & F5 & \text{Class} \\
\hline
\text{Example 1} & t & t & f & f & f & p \\
\text{Example 2} & f & f & t & t & f & p \\
\text{Example 3} & t & f & f & t & f & p \\
\text{Example 4} & t & f & t & f & t & p \\
\text{Example 5} & f & t & f & f & f & n \\
\text{Example 6} & t & t & f & t & t & n \\
\text{Example 7} & f & t & t & t & t & n \\
\end{array}
\]

\[
\begin{align*}
E(F1) &= \frac{4}{7} * I(3,1) + \frac{3}{7} * I(1,2) = \frac{4}{7} * 0.811 + \frac{3}{7} * 0.918 = 0.857 \\
E(F2) &= \frac{4}{7} * I(1,3) + \frac{3}{7} * I(3,0) = \frac{4}{7} * 0.811 + \frac{3}{7} * 0 = 0.463 \\
E(F3) &= \frac{3}{7} * I(2,1) + \frac{4}{7} * I(2,2) = \frac{3}{7} * 0.918 + \frac{4}{7} * 1 = 0.965 \\
E(F4) &= \frac{4}{7} * I(2,2) + \frac{3}{7} * I(1,2) = \frac{4}{7} * 1 + \frac{3}{7} * 0.918 = 0.965 \\
E(F5) &= \frac{3}{7} * I(1,2) + \frac{4}{7} * I(3,1) = \frac{3}{7} * 0.918 + \frac{4}{7} * 0.811 = 0.857 \\
\end{align*}
\]

\[
\begin{align*}
\text{Gain}(F1) &= I(4,3) - E(F1) = 0.128 \\
\text{Gain}(F2) &= I(4,3) - E(F2) = 0.522 \\
\text{Gain}(F3) &= I(4,3) - E(F3) = 0.020 \\
\text{Gain}(F4) &= I(4,3) - E(F4) = 0.020 \\
\text{Gain}(F5) &= I(4,3) - E(F5) = 0.128 \\
\end{align*}
\]

Since \(\text{Gain}(F2)\) is the highest, F2 becomes the root.

Then, choose from \{F1, F3, F4, F5\} to be F2’s f-child.
Since all examples are in class p, it becomes F2’s f-child.

Next, choose from \{F1, F3, F4, F5\} to be F2’s t-child.

<table>
<thead>
<tr>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td>t</td>
<td>f</td>
<td>f</td>
<td>p</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>p</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>t</td>
<td>f</td>
<td>t</td>
<td>p</td>
</tr>
</tbody>
</table>

\[
E(F1) = \frac{2}{4} \times I(1,1) + \frac{2}{4} \times I(0,2) = \frac{2}{4} \times 1 + \frac{2}{4} \times 0 = 0.5
\]
\[
E(F3) = \frac{1}{4} \times I(0,1) + \frac{3}{4} \times I(1,2) = \frac{1}{4} \times 0 + \frac{3}{4} \times 0.918 = 0.6885
\]
\[
E(F4) = \frac{2}{4} \times I(0,2) + \frac{2}{4} \times I(1,1) = \frac{2}{4} \times 0 + \frac{2}{4} \times 1 = 0.5
\]
\[
E(F5) = \frac{2}{4} \times I(0,2) + \frac{2}{4} \times I(1,1) = \frac{2}{4} \times 0 + \frac{2}{4} \times 1 = 0.5
\]

Gain(F1) = I(1,3) - E(F1) = 0.311
Gain(F3) = I(1,3) - E(F3) = 0.1225
Gain(F4) = I(1,3) - E(F4) = 0.311
Gain(F5) = I(1,3) - E(F5) = 0.311

F1, F4, and F5 have the maximum Gain(), we break ties in favor of features with smaller numbers and thus choose F1 to be F2’s t-child.

Then, choose from \{F3, F4, F5\} to be F1’s f-child.

<table>
<thead>
<tr>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>t</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>n</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>f</td>
<td>t</td>
<td>t</td>
<td>n</td>
</tr>
</tbody>
</table>

Therefore, class “n” becomes F1’s f-child.

Then, choose from \{F3, F4, F5\} to be F1’s t-child.
\[ E(F_3) = 0/2 * I(0,0) + 2/2 * I(1,1) = 0/2 * 0 + 2/2 * 1 = 1 \]
\[ E(F_4) = 1/2 * I(0,1) + 1/2 * I(1,0) = 1/2 * 0 + 1/2 * 0 = 0 \]
\[ E(F_5) = 1/2 * I(0,1) + 1/2 * I(1,0) = 1/2 * 0 + 1/2 * 0 = 0 \]

Gain(F_3) = I(1,3) - E(F_3) = 0
Gain(F_4) = I(1,3) - E(F_4) = 1
Gain(F_5) = I(1,3) - E(F_5) = 1

\[ F_4 \text{ and } F_5 \text{ have the highest Gain(). } F_4 \text{ are favored by the tie-breaking scheme and, thus, becomes } F_1 \text{'s t-child.} \]

Next, choose either F_3 or F_5 to be F_4’s f-child.

\[ \begin{array}{cccccc}
F_1 & F_2 & F_3 & F_4 & F_5 & \text{Class} \\
\hline
\text{Example 1} & t & t & f & f & f & p \\
\text{Example 6} & t & t & f & t & t & n
\end{array} \]

Since the only example has class p, “p” becomes F_4’s f-child.

Then, choose either F_3 or F_5 to be F_4’s t-child.

\[ \begin{array}{cccccc}
F_1 & F_2 & F_3 & F_4 & F_5 & \text{Class} \\
\hline
\text{Example 6} & t & t & f & t & t & n
\end{array} \]

For similar reason as before, class n becomes F_4’s t-child.

The final tree below will classify example8 (f,f,t,t,t) as belonging to class p.