**Necessary Formulas:**

1. Entropy, E= –∑ pi log2 pi |;i = 1 to k, where k = number of classes.
2. Average Entropy, ENew = ( ∑ − Vij log2 Vij + ∑ Sj log2 Sj )/N|;

i = 1 to k, where k = number of classes and

j = 1 to n, where n = number of unique values for an attribute and

1. Information Gain, Ig = EStart - ENew

**Iteration 1 (For Selecting the Root Node)**

We have 3 classes. So, The Value of Initial Entropy, EStart will be:

***EStart = – p1 log2 p1 – p2 log2 p2 – p3 log2 p3***

There are 4 instances with classification 1, 5 instances with classification 2 and 15 instances with classification 3. So, p1 = (4/24), p2 = (5/24) and p3 = (15/24).

***EStart = –*** ***(4/24)log2 (4/24)– (5/24)log2 (5/24)– (15/24)log2 (15/24)***

***= 0.4308 + 0.4715 + 0.4238***

***= 1.3261 bits***

Now, we need to calculate ENew for each of the attributes.

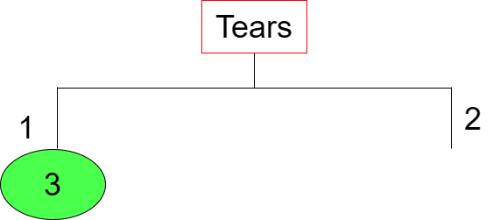
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Frequency Table for Age**   |  |  |  |  | | --- | --- | --- | --- | |  | Age = 1 | Age = 2 | Age = 3 | | Class 1 | 2 | 1 | 1 | | Class 2 | 2 | 2 | 1 | | Class 3 | 4 | 5 | 6 | | Sum | 8 | 8 | 8 | | |  |  |  | | --- | --- | --- | | ENew (Age) | = | (− 2 log2 2 − 1 log2 1 − 1 log2 1  − 2 log2 2 − 2 log2 2 − 1 log2 1  − 4 log2 4 − 5 log2 5 − 6 log2 6  + 8 log2 8 + 8 log2 8 + 8 log2 8) / 24 | |  | = | 1.2867 | |
| **Frequency Table for SpecRx**   |  |  |  | | --- | --- | --- | |  | SpecRx = 1 | SpecRx = 2 | | Class 1 | 3 | 1 | | Class 2 | 2 | 3 | | Class 3 | 7 | 8 | | Sum | 12 | 12 | | |  |  |  | | --- | --- | --- | | ENew (SpecRx) | = | (− 3 log2 3 − 1 log2 1 − 2 log2 2  − 3 log2 3 − 7 log2 7 − 8 log2 8  + 12 log2 12 + 12 log2 12)/24 | |  | = | 1.2866 | |
| **Frequency Table for Astig**   |  |  |  | | --- | --- | --- | |  | Astig = 1 | Astig = 2 | | Class 1 | 0 | 4 | | Class 2 | 5 | 0 | | Class 3 | 7 | 8 | | Sum | 12 | 12 | | |  |  |  | | --- | --- | --- | | ENew (Astig) | = | (− 0 − 4 log2 4 − 5 log2 5 – 0  − 7 log2 7 − 8 log2 8 + 12 log2 12  + 12 log2 12)/24 | |  | = | 0.9491 | |
| **Frequency Table for Tears**   |  |  |  | | --- | --- | --- | |  | Tears = 1 | Tears = 2 | | Class 1 | 0 | 4 | | Class 2 | 0 | 5 | | Class 3 | 12 | 3 | | Sum | 12 | 12 | | |  |  |  | | --- | --- | --- | | ENew (Tears) | = | (− 0 − 4 log2 4 – 0 – 5 log2 5  − 12 log2 12 − 3 log2 3 + 12 log2 12  + 12 log2 12)/24 | |  | = | 0.7773 | |

Ig (Age) = EStart – ENew (Age) = 1.3261 – 1.2867 = 0.0394

Ig (SpecRx) = EStart – ENew (SpecRx) = 1.3261 – 1.2866 = 0.0395

Ig (Astig) = EStart – ENew (Astig) = 1.3261 – 0.9491 = 0.377

Ig (Tears) = EStart – ENew (Tears) = 1.3261 – 0.7773 = **0.5488**



**Iteration 2 (For Branch Tears = 2)**

There are 4 instances with classification 1, 5 instances with classification 2 and 3 instances with classification 3. So, p1 = (4/12), p2 = (5/12) and p3 = (3/12).

***EStart = – (4/12)log2 (4/12)– (5/12)log2 (5/12)– (3/12)log2 (3/12)***

***= 0.5283 + 0.5263 + 0.5***

***= 1.5546 bits***

Now, we need to calculate ENew for each of the attributes.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Frequency Table for Age**   |  |  |  |  | | --- | --- | --- | --- | |  | Age = 1 | Age = 2 | Age = 3 | | Class 1 | 2 | 1 | 1 | | Class 2 | 2 | 2 | 1 | | Class 3 | 0 | 1 | 2 | | Sum | 4 | 4 | 4 | | |  |  |  | | --- | --- | --- | | ENew (Age) | = | (– 2 log2 2 – 1 log2 1 – 1 log2 1 – 2 log2 2 – 2 log2 2 – 1 log2 1 – 0 – 1 log2 1 – 2 log2 2 + 4 log2 4 + 4 log2 4 + 4 log2 4)/12 | |  | = | 1.3333 | |
| **Frequency Table for SpecRx**   |  |  |  | | --- | --- | --- | |  | SpecRx = 1 | SpecRx = 2 | | Class 1 | 3 | 1 | | Class 2 | 2 | 3 | | Class 3 | 1 | 2 | | Sum | 6 | 6 | | |  |  |  | | --- | --- | --- | | ENew (SpecRx) | = | (– 3 log2 3 – 1 log2 1 – 2 log2 2 – 3 log2 3 – 1 log2 1 – 2 log2 2 + 6 log2 6 + 6 log2 6) / 12 | |  | = | 1.4592 | |
| **Frequency Table for Astig**   |  |  |  | | --- | --- | --- | |  | Astig = 1 | Astig = 2 | | Class 1 | 0 | 4 | | Class 2 | 5 | 0 | | Class 3 | 1 | 2 | | Sum | 6 | 6 | | |  |  |  | | --- | --- | --- | | ENew (Astig) | = | (0 – 4 log2 4 – 5 log2 5 – 0 – 1 log2 1 – 2 log2 2 + 6 log2 6 + 6 log2 6)/12 | |  | = | 0.7842 | |

Ig (Age) = EStart – ENew (Age) = 1.5546 – 1.3333 = 0.2213

Ig (SpecRx) = EStart – ENew (SpecRx) = 1.5546 – 1.4592 = 0.0954

Ig (Astig) = EStart – ENew (Astig) = 1.5546 – 0.7842 = **0.7704**



**Iteration 3 (For Branch Astig = 1)**

There are 5 instances with classification 2 and 1 instance with classification 3. So, p1 = (5/6) and p2 = (1/6).

***EStart = – (5/6)log2 (5/6)– (1/6)log2 (1/6)***

***= 0.2192 + 0.4308***

***= 0.65 bits***

Now, we need to calculate ENew for each of the attributes.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Frequency Table for Age**   |  |  |  |  | | --- | --- | --- | --- | |  | Age = 1 | Age = 2 | Age = 3 | | Class 1 | 0 | 0 | 0 | | Class 2 | 2 | 2 | 1 | | Class 3 | 0 | 0 | 1 | | Sum | 2 | 2 | 2 | | |  |  |  | | --- | --- | --- | | ENew (Age) | = | (0 – 0 – 0 – 2 log2 2 – 2 log2 2 – 1 log2 1 – 0 – 0 – 1 log2 1 + 2 log2 2 + 2 log2 2 + 2 log2 2)/6 | |  | = | 0.3333 | |
| **Frequency Table for SpecRx**   |  |  |  | | --- | --- | --- | |  | SpecRx = 1 | SpecRx = 2 | | Class 1 | 0 | 0 | | Class 2 | 2 | 3 | | Class 3 | 1 | 0 | | Sum | 3 | 3 | | |  |  |  | | --- | --- | --- | | ENew (SpecRx) | = | (0 – 0 – 2 log2 2 – 3 log2 3 – 1 log2 1 – 0 + 3 log2 3 + 3 log2 3)/6 | |  | = | 0.4592 | |

Ig (Age) = EStart – ENew (Age) = 0.6500 – 0.3333 = **0.3167**

Ig (SpecRx) = EStart – ENew (SpecRx) = 0.6500 – 0.4592 = 0.1908

A diagram of a number of people

AI-generated content may be incorrect.

**Iteration 4 (For Branch Astig = 2)**

There are 4 instances with classification 1 and 2 instance with classification 3. So, p1 = (4/6) and p2 = (2/6).

***EStart = – (4/6)log2 (4/6)– (2/6)log2 (2/6)***

***= 0.3900 + 0.5283***

***= 0.9183 bits***

Now, we need to calculate ENew for each of the attributes.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Frequency Table for Age**   |  |  |  |  | | --- | --- | --- | --- | |  | Age = 1 | Age = 2 | Age = 3 | | Class 1 | 2 | 1 | 1 | | Class 2 | 0 | 0 | 0 | | Class 3 | 0 | 1 | 1 | | Sum | 2 | 2 | 2 | | |  |  |  | | --- | --- | --- | | ENew (Age) | = | (– 2 log2 2 – 1 log2 1 – 1 log2 1 – 0 – 0 –0 – 0 – 1 log2 1 – 1 log2 1 + 2 log2 2 + 2 log2 2 + 2 log2 2)/6 | |  | = | 0.6667 | |
| **Frequency Table for SpecRx**   |  |  |  | | --- | --- | --- | |  | SpecRx = 1 | SpecRx = 2 | | Class 1 | 3 | 1 | | Class 2 | 0 | 0 | | Class 3 | 0 | 2 | | Sum | 3 | 3 | | |  |  |  | | --- | --- | --- | | ENew (SpecRx) | = | (– 3 log2 3 – 1 log2 1 – 0 – 0 – 0 – 2 log2 2 + 3 log2 3 + 3 log2 3)/6 | |  | = | 0.4592 | |

Ig (Age) = EStart – ENew (Age) = 0.9183 – 0.6667 = 0.2516

Ig (SpecRx) = EStart – ENew (SpecRx) = 0.9183 – 0.4592 = **0.4591**

A black background with green circles and white text

AI-generated content may be incorrect.