Iterate on all humidity values and separate dataset into two parts as instances less than or equal to current value, and instances greater than the current value. We would calculate the gain or gain ratio for every step. The value which maximizes the gain would be the threshold.

Check 65 as a threshold for humidity

Entropy(Decision|Humidity<=65) = – p(No) . log2p(No) – p(Yes) . log2p(Yes) = -(0/1).log2(0/1) – (1/1).log2(1/1) = 0

Entropy(Decision|Humidity>65) = -(5/13).log2(5/13) – (8/13).log2(8/13) =0.530 + 0.431 = 0.961

Gain(Decision, Humidity<> 65) = 0.940 – (1/14).0 – (13/14).(0.961) = 0.048

SplitInfo(Decision, Humidity<> 65) = -(1/14).log2(1/14) -(13/14).log2(13/14) = 0.371

GainRatio(Decision, Humidity<> 65) = 0.126

Check 70 as a threshold for humidity

Entropy(Decision|Humidity<=70) = – (1/4).log2(1/4) – (3/4).log2(3/4) = 0.811

Entropy(Decision|Humidity>70) =  – (4/10).log2(4/10) – (6/10).log2(6/10) = 0.970

Gain(Decision, Humidity<> 70) = 0.940 – (4/14).(0.811) – (10/14).(0.970) = 0.940 – 0.231 – 0.692 = 0.014

SplitInfo(Decision, Humidity<> 70) = -(4/14).log2(4/14) -(10/14).log2(10/14) = 0.863

GainRatio(Decision, Humidity<> 70) = 0.016

Check 75 as a threshold for humidity

Entropy(Decision|Humidity<=75) = – (1/5).log2(1/5) – (4/5).log2(4/5) = 0.721

Entropy(Decision|Humidity>75) = – (4/9).log2(4/9) – (5/9).log2(5/9) = 0.991

Gain(Decision, Humidity<> 75) = 0.940 – (5/14).(0.721) – (9/14).(0.991) = 0.940 – 0.2575 – 0.637 = 0.045

SplitInfo(Decision, Humidity<> 75) = -(5/14).log2(4/14) -(9/14).log2(10/14) = 0.940

GainRatio(Decision, Humidity<> 75) = 0.047

Gain(Decision, Outlook <> 78) =0.090, GainRatio(Decision, Humidity<> 78) =0.090

Gain(Decision, Outlook <> 80) = 0.101, GainRatio(Decision, Humidity<> 80) = 0.107

Gain(Decision, Outlook <> 85) = 0.024, GainRatio(Decision, Humidity<> 85) = 0.027

Gain(Decision, Outlook <> 90) = 0.010, GainRatio(Decision, Humidity<> 90) = 0.016

As seen, gain maximizes when threshold is equal to 80 for humidity. This means that we need to compare other nominal attributes and comparison of humidity to 80 to create a branch in our tree.

Let’s summarize calculated gain and gain ratios. Outlook attribute comes with both maximized gain and gain ratio. This means that we need to put outlook decision in root of decision tree.