

We implement Decision Tree with ID3 Algorithm. The following formulae are used:

Entropy (D) = - (p₁ * log₂p₁ + p₂ * log₂p₂)

InfoGain (featured)=Entropy(D)–Entropy(featured)

1st Iteration:

Entropy Calculation

Hired	
Yes	No
17	47

Entropy (Hired) = Entropy (17,47)

= Entropy (0.266, 0.734)

= - (0.266 log₂ 0.266) - (0.734 log₂ 0.734)

= 0.836

Considering Resume Compatibility as primary splitting node,

		Hired		
		Yes	No	
Resume Compatibility	Yes	17	15	32
	No	0	32	32
				64

Entropy (Hired, Resume Compatibility) = P(Yes)*E(17,15) + P(No)*E(0,32)

= 0.5 * 0.997 + 0.5 * 0

= 0.4985

Information Gain = 0.836 – 0.4985

=0.3375

Similarly,

		Hired		
		Yes	No	
Resume Compatibility	Yes	17	15	32
	No	0	32	32
TOTAL:				64
Information Gain:				0.3375

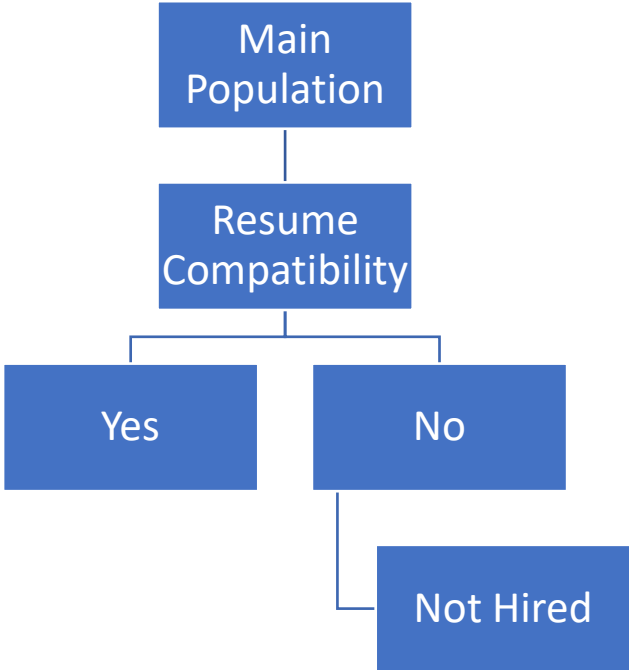
		Hired		
		Yes	No	
Premium Institute	Yes	12	20	32
	No	5	27	32
TOTAL:				64
Information Gain:				0.046
		Hired		
		Yes	No	
Work Ex	>12	11	21	32
	<12	6	26	32
TOTAL:				64
Information Gain:				0.0237

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	10	22	32
	High	7	25	32
TOTAL:				64
Information Gain:				0.009

		Hired		
		Yes	No	
Interview Performance	Impressive	12	20	32
	Mediocre	5	27	32
TOTAL:				64
Information Gain:				0.046

		Hired		
		Yes	No	
Recommendation	Yes	11	21	32
	No	6	24	32
TOTAL:				64
Information Gain:				0.0237

Therefore, Resume Compatibility is the first Decision Node.



2nd Iteration:

Hired	
Yes	No
17	15

Entropy (Hired) = Entropy (17,15)

= Entropy (0.531, 0.469)

= 0.997

		Hired		
		Yes	No	
Premium Institute	Yes	12	4	16
	No	5	11	16
TOTAL:				32
Information Gain:				0.143

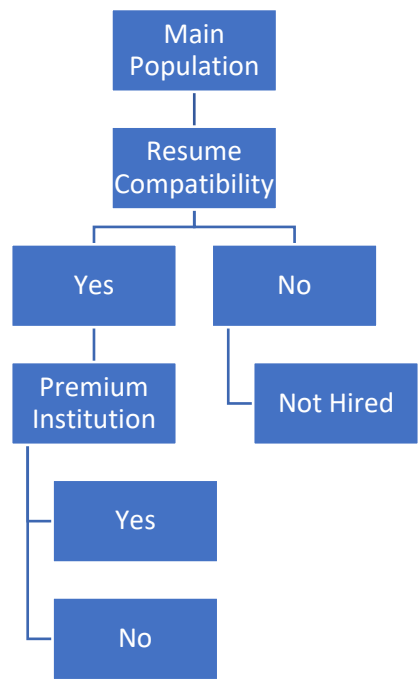
		Hired		
		Yes	No	
Work Ex	>12	11	5	16
	<12	6	10	16
TOTAL:				32
Information Gain:				0.071

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	10	6	16
	High	7	9	16
TOTAL:				32
Information Gain:				0.025

		Hired		
		Yes	No	
Interview Performance	Impressive	12	4	16
	Mediocre	5	11	16
TOTAL:				32
Information Gain:				0.143

		Hired		
		Yes	No	
Recommendation	Yes	11	5	16
	No	6	10	16
TOTAL:				32
Information Gain:				0.071

As you can see, we can take either Interview Performance or Premium Institute factor as both of them provide the same maximum value of Information Gain. We choose Premium Institute.



3rd Iteration:

For the Yes branch,

Hired	
Yes	No
12	4

Entropy (Hired) = Entropy (12,4)

= Entropy (0.75, 0.25)

= 0.811

		Hired		
		Yes	No	
Work Ex	>12	7	1	8
	<12	5	3	8
TOTAL:				16
Information Gain:				0.062

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	6	2	8
	High	6	2	8
TOTAL:				16
Information Gain:				0

		Hired		
		Yes	No	
Interview Performance	Impressive	8	0	8
	Mediocre	4	4	8
TOTAL:				16
Information Gain:				0.311

		Hired		
		Yes	No	
Recommendation	Yes	7	1	8
	No	5	3	8
TOTAL:				16
Information Gain:				0.062

Thus, Interview Performance is selected.

For the No branch,

Hired	
Yes	No
5	11

Entropy (Hired) = Entropy (5,11)

= Entropy (0.3125, 0.6875)

= 0.896

		Hired		
		Yes	No	
Work Ex	>12	4	4	8
	<12	1	7	8
TOTAL:				16
Information Gain:				0.124

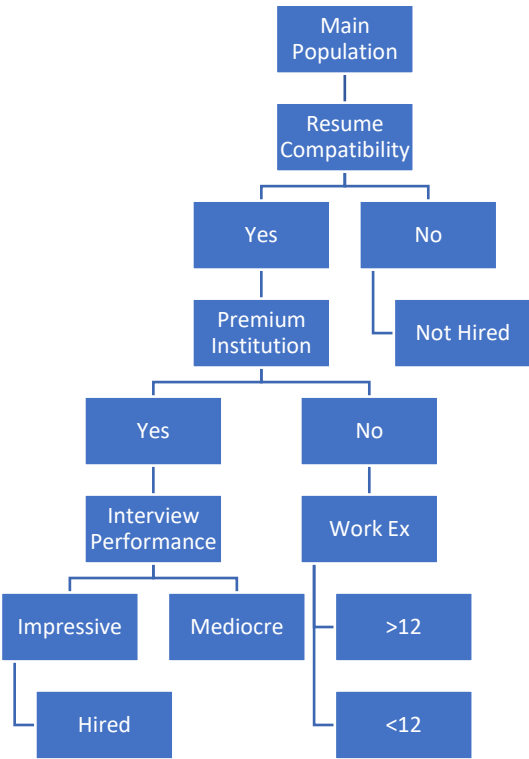
		Hired		
		Yes	No	
Expected Remuneration	Reasonable	4	4	8
	High	1	7	8
TOTAL:				16
Information Gain:				0.124

		Hired		
		Yes	No	
Interview Performance	Impressive	4	4	8
	Mediocre	1	7	8
TOTAL:				16
Information Gain:				0.124

		Hired		
		Yes	No	
Recommendation	Yes	4	4	8
	No	1	7	8
TOTAL:				16
Information Gain:				0.124

As all give same Information Gain, we arbitrarily select Work Ex.

Thus, the augmented tree is as follows:



4th Iteration:

For Mediocre under Interview Performance,

Hired	
Yes	No
4	4

Entropy (Hired) = Entropy (4,4)

= Entropy (0.5, 0.5)

= 1

		Hired		
		Yes	No	
Work Ex	>12	3	1	4
	<12	1	3	4
TOTAL:				8
Information Gain:				0.189

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	2	2	4
	High	2	2	4
TOTAL:				8
Information Gain:				0

		Hired		
		Yes	No	
Recommendation	Yes	3	1	4
	No	1	3	4
TOTAL:				8
Information Gain:				0.189

We arbitrarily select Recommendation.

For >12 under Work Ex,

Hired	
Yes	No
4	4

Entropy (Hired) = Entropy (4,4)

= Entropy (0.5, 0.5)

= 1

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	3	1	4
	High	1	3	4
TOTAL:				8
Information Gain:				0.189

		Hired		
		Yes	No	
Interview Performance	Impressive	3	1	4
	Mediocre	1	3	4
TOTAL:				8
Information Gain:				0.189

		Hired		
		Yes	No	
Recommendation	Yes	3	1	4
	No	1	3	4
TOTAL:				8
Information Gain:				0.189

We arbitrarily select Recommendation.

For <12 under Work Ex,

Hired	
Yes	No
1	7

Entropy (Hired) = Entropy (1,7)

= Entropy (0.125, 0.875)

= 0.544

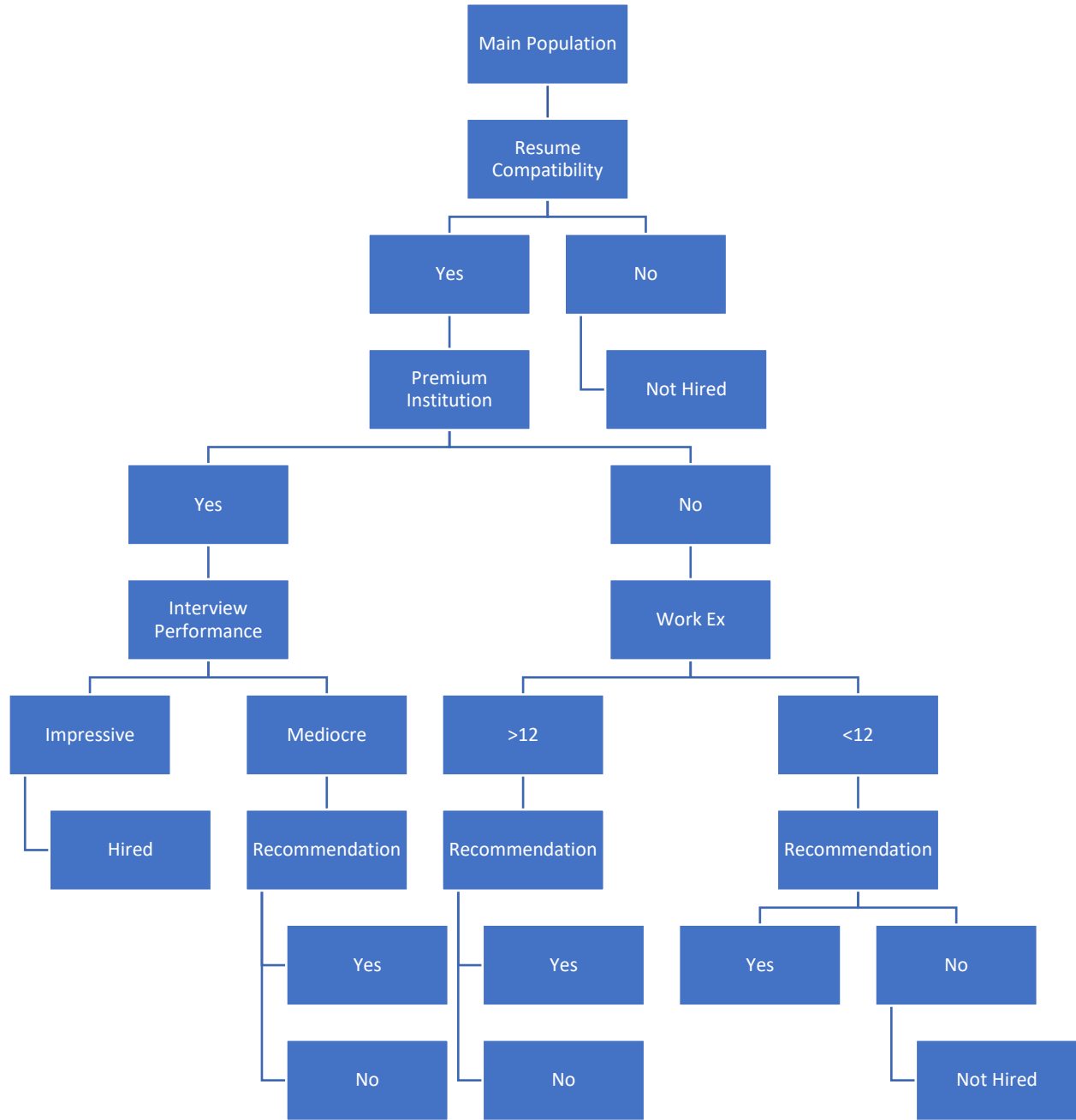
		Hired		
		Yes	No	
Expected Remuneration	Reasonable	1	3	4
	High	0	4	4
TOTAL:				8
Information Gain:				0.594

		Hired		
		Yes	No	
Interview Performance	Impressive	1	3	4
	Mediocre	0	4	4
TOTAL:				8
Information Gain:				0.594

		Hired		
		Yes	No	
Recommendation	Yes	1	3	4
	No	0	4	4
TOTAL:				8
Information Gain:				0.594

We arbitrarily select Recommendation.

Thus, the Augmented Tree becomes:



5th Iteration:

For the first Yes branch,

Hired	
Yes	No
3	1

Entropy (Hired) = Entropy (3,1)

= Entropy (0.75, 0.25)

= 0.811

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	1	1	2
	High	2	0	2
TOTAL:				4
Information Gain:				0.311

		Hired		
		Yes	No	
Work Ex	>12	2	0	2
	<12	1	1	2
TOTAL:				4
Information Gain:				0.311

We arbitrarily select Work Ex.

For the first No branch,

Hired				
Yes		No		
1		3		

Entropy (Hired) = Entropy (1,3)

= Entropy (0.25, 0.75)

= 0.811

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	1	1	2
	High	0	2	2
TOTAL:				4
Information Gain:				0.311

		Hired		
		Yes	No	
Work Ex	>12	1	1	2
	<12	0	2	2
TOTAL:				4
Information Gain:				0.311

We arbitrarily select Work Ex.

For the second Yes branch,

Hired				
Yes		No		
3		1		

Entropy (Hired) = Entropy (3,1)

= Entropy (0.75, 0.25)

= 0.811

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	2	0	2
	High	1	1	2
TOTAL:				4
Information Gain:				0.311

		Hired		
		Yes	No	
Interview Performance	Impressive	2	0	2
	Mediocre	1	1	2
TOTAL:				4
Information Gain:				0.311

We arbitrarily select Interview Performance.

For the second No branch,

Hired				
Yes		No		
1		3		

Entropy (Hired) = Entropy (1,3)

= Entropy (0.25, 0.75)

= 0.811

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	1	1	2
	High	0	2	2
TOTAL:				4
Information Gain:				0.311

		Hired		
		Yes	No	
Interview Performance	Impressive	1	1	2
	Mediocre	0	2	2
TOTAL:				4
Information Gain:				0.311

We arbitrarily select Interview Performance.

For the third Yes branch,

Hired	
Yes	No
1	3

Entropy (Hired) = Entropy (1,3)

= Entropy (0.25, 0.75)

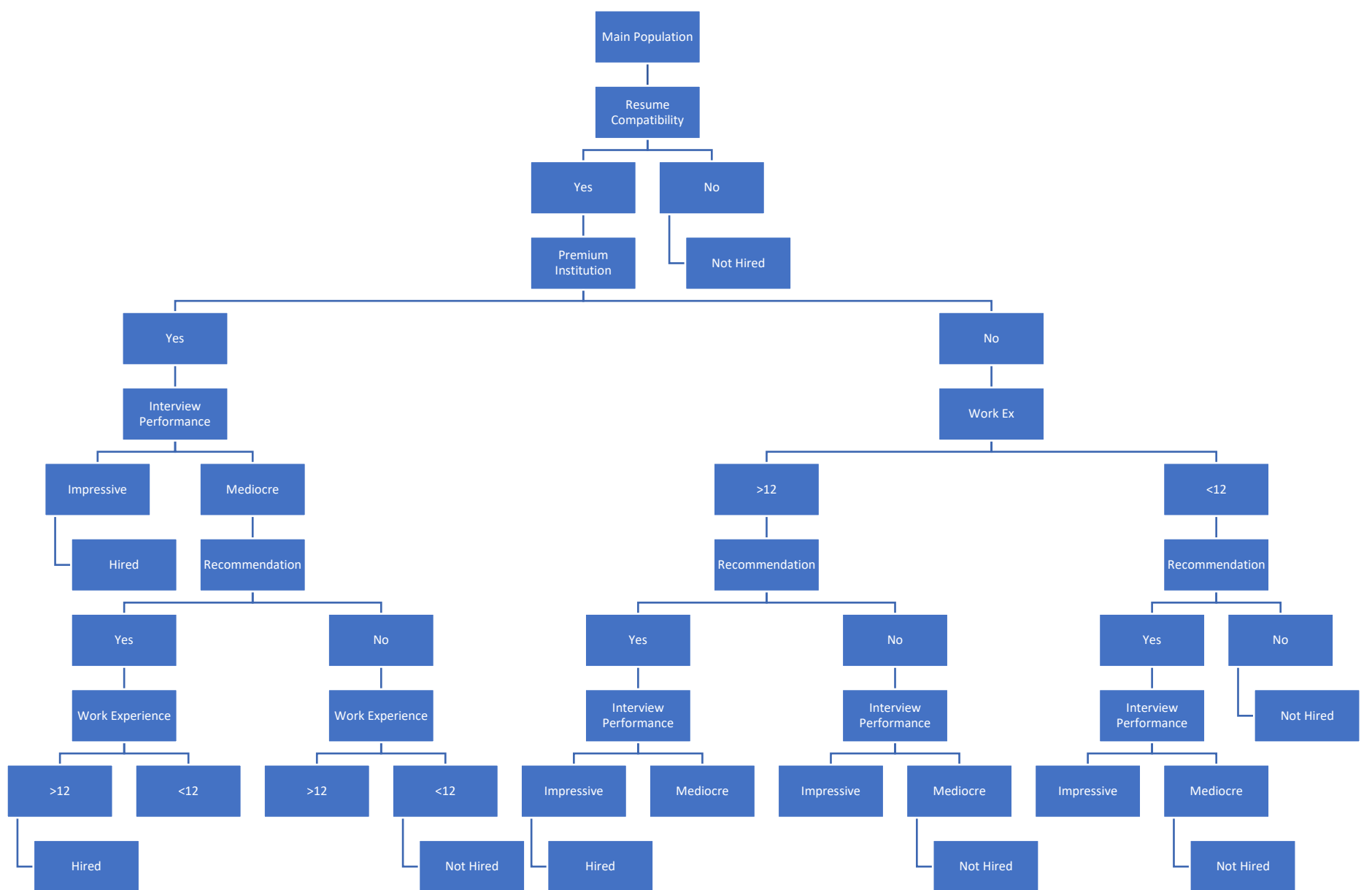
= 0.811

		Hired		
		Yes	No	
Expected Remuneration	Reasonable	1	1	2
	High	0	2	2
TOTAL:				4
Information Gain:				0.311

		Hired		
		Yes	No	
Interview Performance	Impressive	1	1	2
	Mediocre	0	2	2
TOTAL:				4
Information Gain:				0.311

We arbitrarily select Interview Performance.

The present Augmented Tree:



6th/Final Iteration:

Now, we simply need to augment whatever criteria option we have left for leaf nodes in the tree above which are not final nodes, i.e. they don't tell us about the classification in terms of hiring.

Thus, the tree becomes as such:

