

Practice Problem on Regression Tree

When you're looking for a community to live in, there are a lot of factors you might consider, e.g., home affordability, proximity to work, and quality of the nearby natural environment. Livability is a measurement of how attractive a neighborhood, city, and/or region is for you based on a variety of factors.

Develop a non-linear machine learning model to predict the livability scores of different cities of this world considering the following dataset. Consider the expansion threshold as 2.

	<i>City</i>	<i>Safeness score</i>	<i>Environment</i>	<i>Affordability</i>	<i>Livability Score?</i>
	<i>Tokyo</i>	88	Excellent	Expensive	80
	<i>Seoul</i>	86	Good	Affordable	78
	<i>Singapore</i>	84	Good	Expensive	75
	<i>Delhi</i>	50	Unhealthy	Affordable	50
	<i>New York</i>	70	Good	Expensive	79
	<i>Karachi</i>	40	Unhealthy	Affordable	47
	<i>Amsterdam</i>	90	Excellent	Affordable	86

Infer the livability score of Dhaka city based on the provided feature values:

<i>City</i>	<i>Safeness score</i>	<i>Environment</i>	<i>Affordability</i>	<i>Livability Score?</i>
<i>Dhaka</i>	52	Unhealthy	Expensive	?

City	Safeness score	Environment	Affordability	Livability Score?
Tokyo	88	Excellent	Expensive	80
Seoul	86	Good	Affordable	78
Singapore	84	Good	Expensive	75
Delhi	50	Unhealthy	Affordable	50
New York	70	Good	Expensive	79
Karachi	40	Unhealthy	Affordable	47
Amsterdam	90	Excellent	Affordable	86

	Safeness score	Environment	Affordability	Livability Score?
Karachi	40	Unhealthy	Affordable	47
Delhi	50	Unhealthy	Affordable	50
New York	70	Good	Expensive	79
Singapore	84	Good	Expensive	75
Seoul	86	Good	Affordable	78
Tokyo	88	Excellent	Expensive	80
Amsterdam	90	Excellent	Affordable	86

<i>Safeness score</i>	Average of consecutive two safeness scores
40	
50	45
70	60
84	77
86	85
88	87
90	89

	Safeness score	Livability Score?		Average Livability Score of Left side instances	Average Livability Score of Right side instances	Sum of square residuals (SSR)
1	40	47	Considering safeness score 45 as the separator	47	74.66666667	795.3333333
2	50	50	Considering safeness score 60 as the separator	48.5	79.6	69.7
3	70	79	Considering safeness score 77 as the separator	58.66666667	79.75	689.4166667
4	84	75	Considering safeness score 85 as the separator	62.75	81.33333333	859.4166667
5	86	78	Considering safeness score 87 as the separator	65.8	83	1028.8
6	88	80	Considering safeness score 89 as the separator	68.16666667	86	1178.833333
7	90	86				
Minimum SSR of Safeness score=69.7						69.7

<i>Environment</i>	<i>Livability Score?</i>		Average Livability Score of Left side instances	Average Livability Score of Right side instances	Sum of square residuals (SSR)
Excellent	80	Considering 'Environment=Excellent' as the separator	83	65.8	1028.8
Good	78	Considering 'Environment=Good' as the separator	77.33333333	65.75	1221.416667
Good	75	Considering 'Environment=Unhealthy' as the separator	48.5	79.6	69.7
Unhealthy	50				
Good	79				
Unhealthy	47				
Excellent	86				
Minimum SSR for Environment=69.7					69.7

<i>Affordability</i>	<i>Livability Score?</i>
Expensive	80
Affordable	78
Expensive	75
Affordable	50
Expensive	79
Affordable	47
Affordable	86

Considering
'Affordability=Expensive' as the separator

Average Livability
Score of Left side
instances

78

Average Livability
Score of Right side
instances

65.25

Sum of square residuals
(SSR)

1172.75

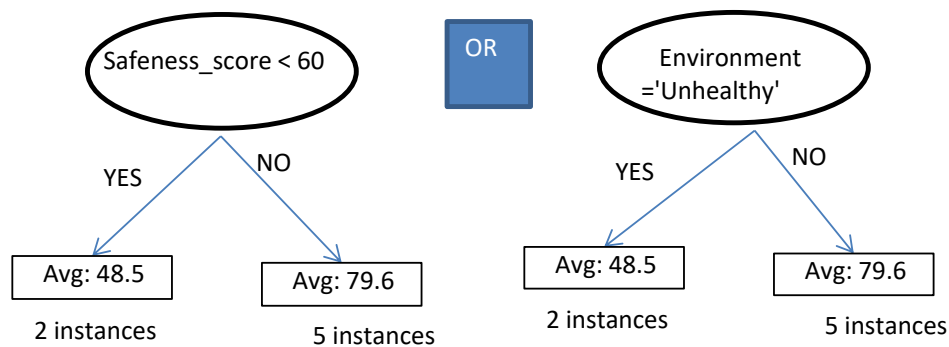
Minimum SSR for Affordability=1172.75

1172.75

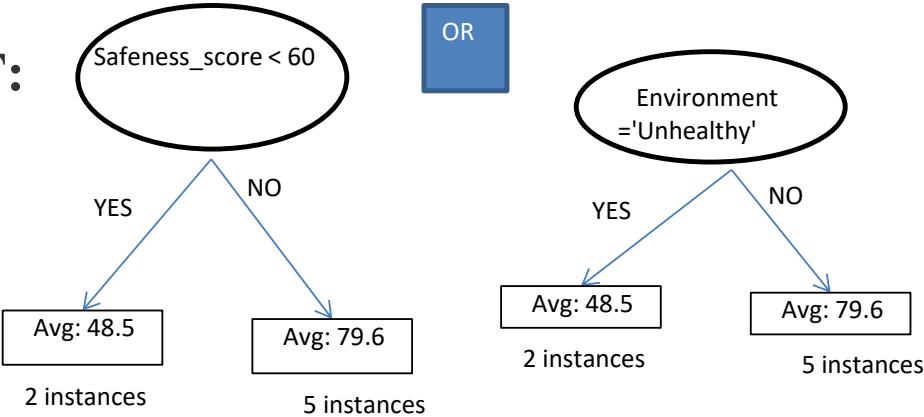
Minimum SSR of Safeness score=69.7

Minimum SSR for Environment=69.7

Minimum SSR for Affordability=1172.75



FOR the
Following RT:



	<i>Safeness score</i>	<i>Environment</i>	<i>Affordability</i>	<i>Livability Score?</i>
<i>New York</i>	70	Good	Expensive	79
<i>Singapore</i>	84	Good	Expensive	75
<i>Seoul</i>	86	Good	Affordable	78
<i>Tokyo</i>	88	Excellent	Expensive	80
<i>Amsterdam</i>	90	Excellent	Affordable	86

	<i>Safeness score</i>	<i>Livability Score?</i>		Average Livability Score of Left side instances	Average Livability Score of Right side instances	Sum of square residuals (SSR)
3	70	79	Considering safeness score 77 as the separator	79	79.75	64.75
4	84	75	Considering safeness score 85 as the separator	77	81.33333333	42.66666667
5	86	78	Considering safeness score 87 as the separator	77.33333333	83	26.66666667
6	88	80	Considering safeness score 89 as the separator	78	86	14
7	90	86				14
Minimum SSR of Safeness score=14						

	<i>Environment</i>	<i>Livability Score?</i>		Average Livability Score of Left side instances	Average Livability Score of Right side instances	Sum of square residuals (SSR)
	Good	79	Considering 'Environment=Excel lent' as the separator	83	77.33333333	26.66666667
	Good	75				
	Good	78				
	Excellent	80				
	Excellent	86				
Minimum SSR for Environment=26.67						26.66666667

<i>Affordability</i>	<i>Livability Score?</i>
Expensive	79
Expensive	75
Affordable	78
Expensive	80
Affordable	86

Considering
'Affordability=Expensive' as the
separator

**Average Livability
Score of Left side
instances**

78

**Average Livability
Score of Right side
instances**

82

**Sum of square residuals
(SSR)**

46

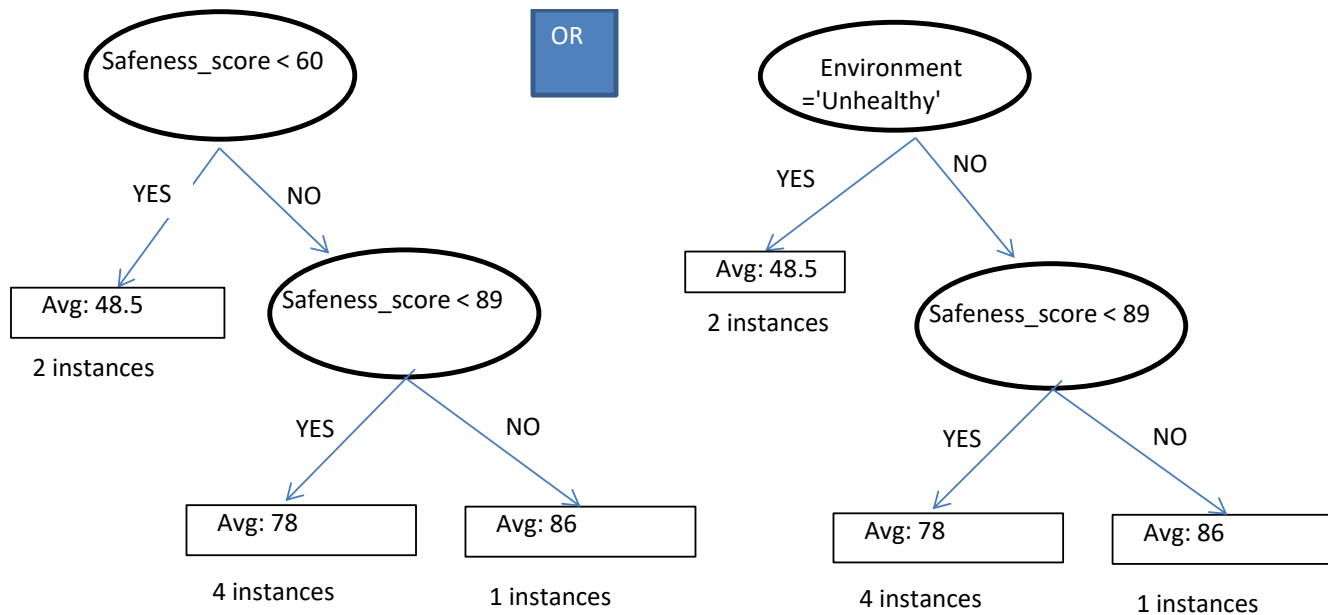
Minimum SSR for Affordability=46

46

Minimum SSR of Safeness score=14

Minimum SSR for Environment=26.67

Minimum SSR for Affordability=46



	<i>Safeness score</i>	<i>Environment</i>	<i>Affordability</i>	<i>Livability Score?</i>
<i>New York</i>	70	Good	Expensive	79
<i>Singapore</i>	84	Good	Expensive	75
<i>Seoul</i>	86	Good	Affordable	78
<i>Tokyo</i>	88	Excellent	Expensive	80

	<i>Safeness score</i>	<i>Livability Score?</i>		Average Livability Score of Left side instances	Average Livability Score of Right side instances	Sum of square residuals (SSR)
3	70	79	Considering safeness score 77 as the separator	79	77.66666667	12.66666667
4	84	75	Considering safeness score 85 as the separator	77	79	10
5	86	78	Considering safeness score 87 as the separator	77.33333333	80	8.66666667
6	88	80				
			Minimum SSR of Safeness score=8.67			8.66666667

	<i>Environment</i>	<i>Livability Score?</i>		Average Livability Score of Left side instances	Average Livability Score of Right side instances	Sum of square residuals (SSR)
	Good	79	Considering 'Environment=Excel lent' as the separator	80	77.33333333	8.66666667
	Good	75				
	Good	78				
	Excellent	80				
			Minimum SSR for Environment=8.67			8.66666667

Affordability	Livability Score?
Expensive	79
Expensive	75
Affordable	78
Expensive	80

Considering
'Affordability=Expensive' as
the separator

Average Livability
Score of Left side
instances

78

Average Livability
Score of Right side
instances

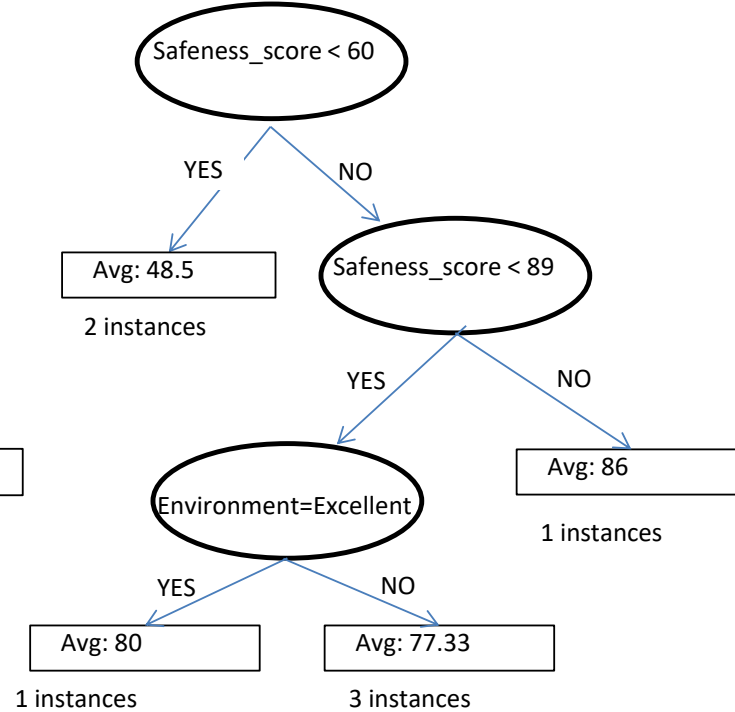
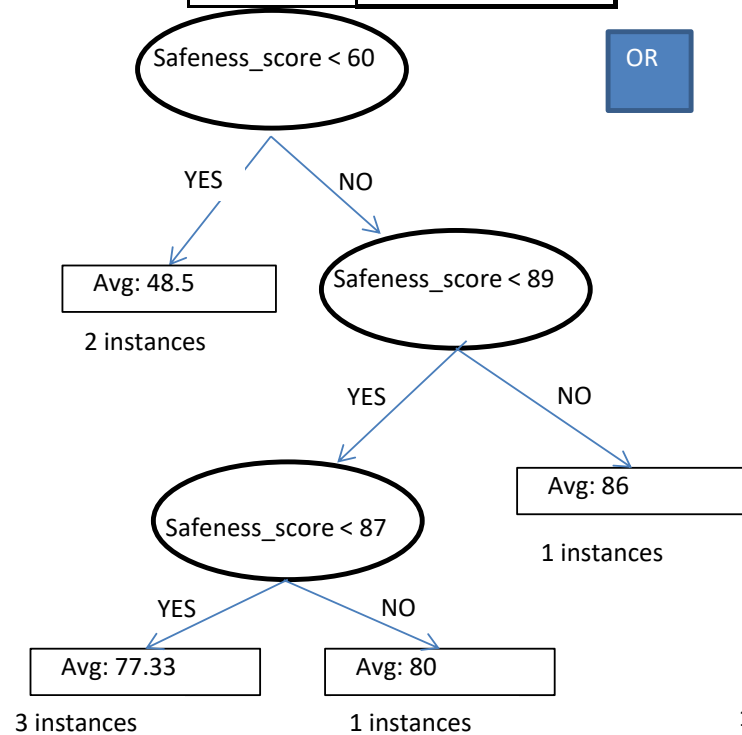
78

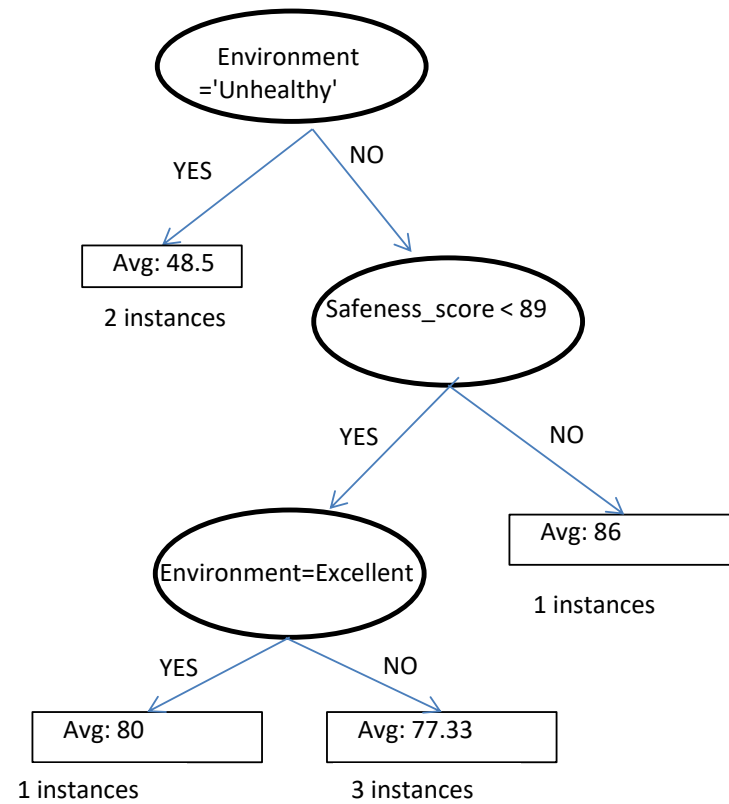
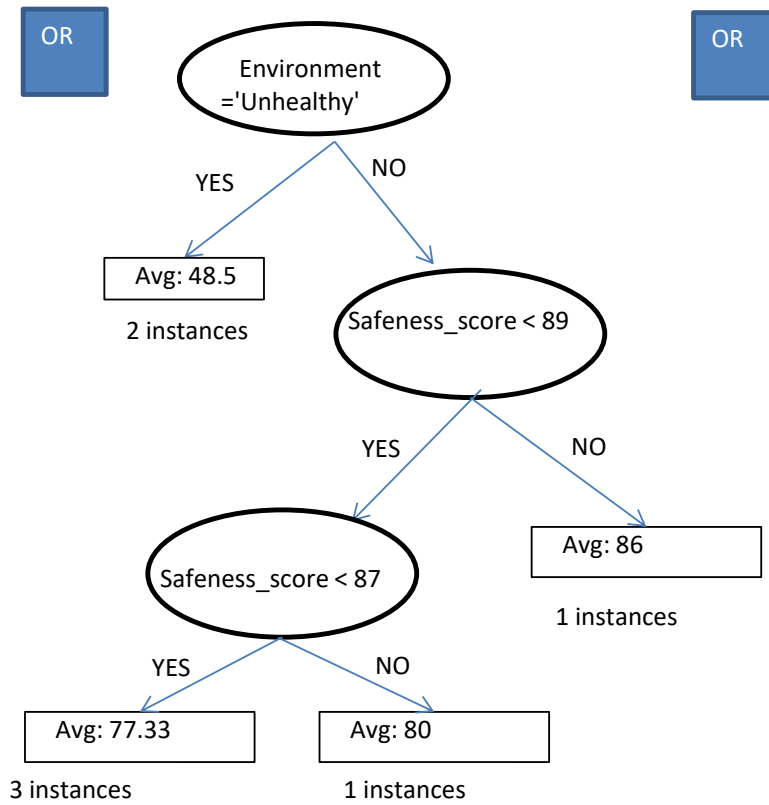
Sum of square residuals
(SSR)

14

Minimum SSR for Affordability=14

14





	<i>Safeness score</i>	<i>Environment</i>	<i>Affordability</i>	<i>Livability Score?</i>
<i>New York</i>	70	Good	Expensive	79
<i>Singapore</i>	84	Good	Expensive	75
<i>Seoul</i>	86	Good	Affordable	78

	Safeness score	Livability Score?	Average Livability Score of Left side instances	Average Livability Score of Right side instances	Sum of square residuals (SSR)
3	70	79	79	76.5	4.5
4	84	75	77	78	8
5	86	78			
			Minimum SSR of Safeness score=4.5		

Environment	Livability Score?	Average Livability Score of Left side instances	Average Livability Score of Right side instances	Sum of square residuals (SSR)
Good	79	77.33333333	0	8.666666667
Good	75			
Good	78			
		Minimum SSR for Environment=8.67		8.666666667

<i>Affordability</i>	<i>Livability Score?</i>
Expensive	79
Expensive	75
Affordable	78

Considering
'Affordability=Expensive' as the
separator

**Average Livability
Score of Left side
instances**

77

**Average Livability
Score of Right side
instances**

78

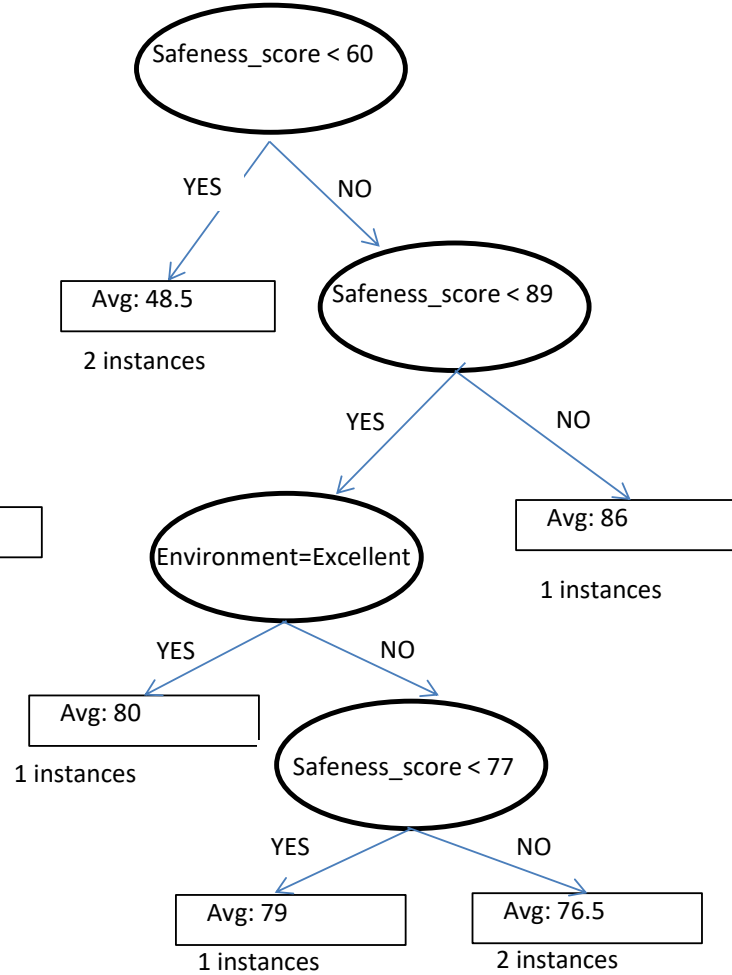
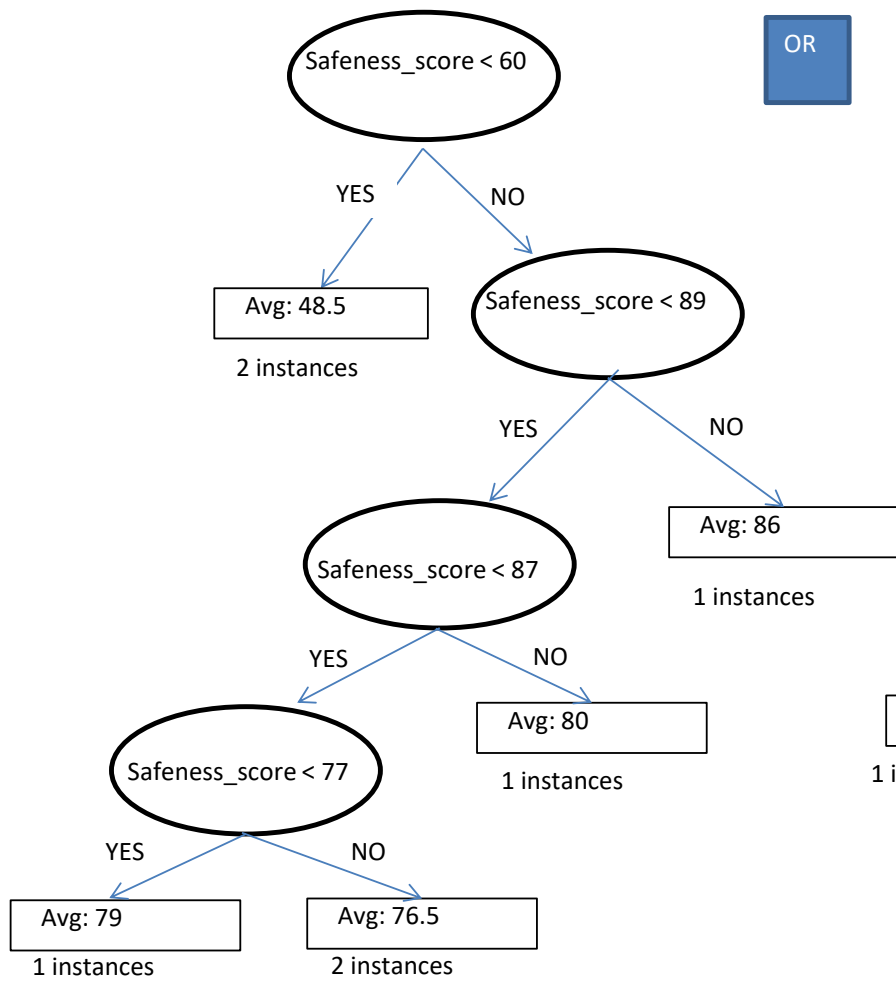
**Sum of square residuals
(SSR)**

8

Minimum SSR for Affordability=14

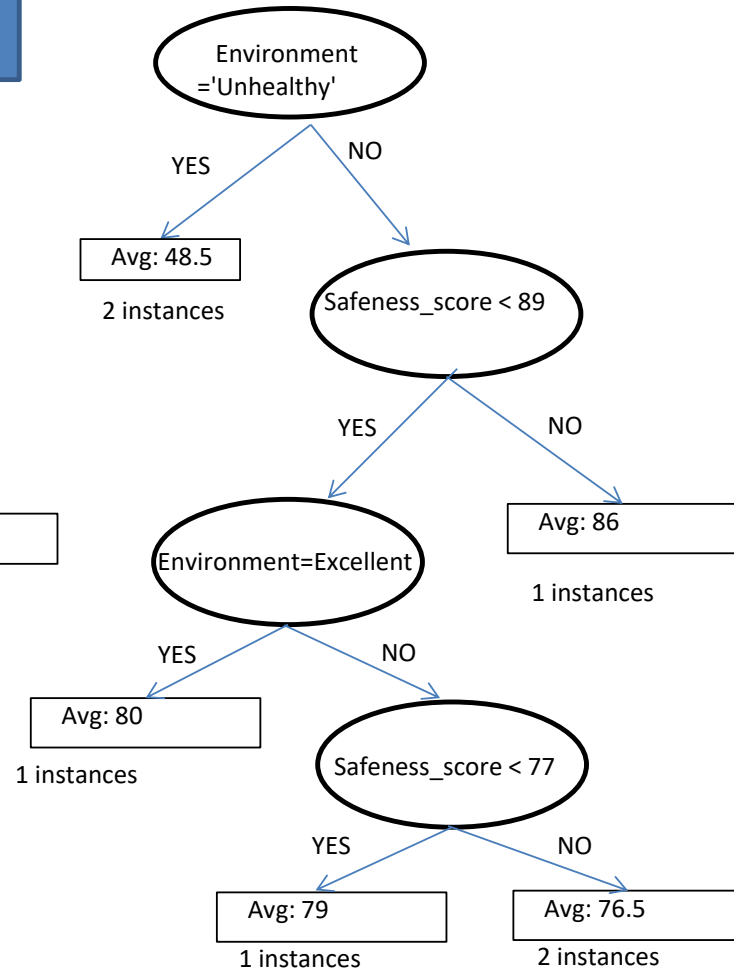
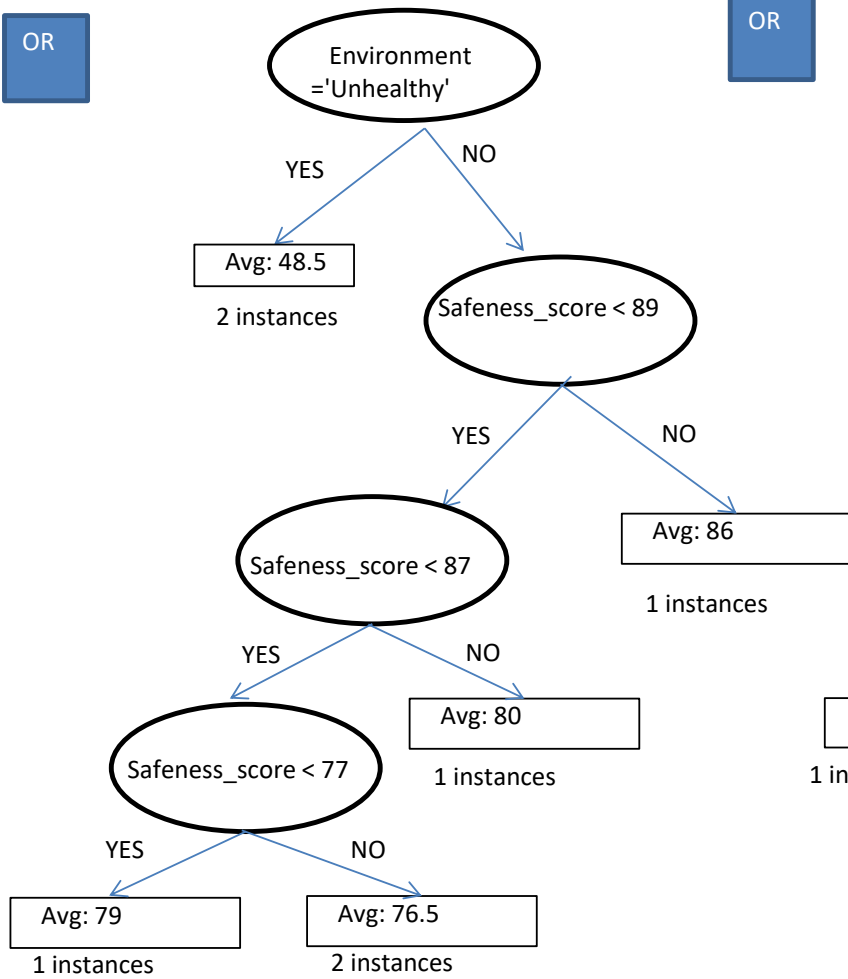
8

OR



OR

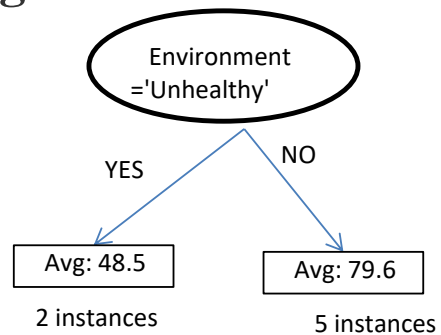
OR



Query:

	<i>Safeness score</i>	<i>Environment</i>	<i>Affordability</i>	<i>Livability Score?</i>
<i>Dhaka</i>	52	Unhealthy	Expensive	48.5

**FOR the
Following RT:**



	<i>Safeness score</i>	<i>Environment</i>	<i>Affordability</i>	<i>Livability Score?</i>
<i>New York</i>	70	Good	Expensive	79
<i>Singapore</i>	84	Good	Expensive	75
<i>Seoul</i>	86	Good	Affordable	78
<i>Tokyo</i>	88	Excellent	Expensive	80
<i>Amsterdam</i>	90	Excellent	Affordable	86