Decision Tree using GINI Index

Main Dataset

	Weekened	Weather	Parents	Money	Decision
1	W1	Sunny	Yes	Rich	Cinema
2	W2	Sunny	No	Rich	Tennis
3	W3	Windy	Yes	Rich	Cinema
4	W4	Rainy	Yes	Poor	Cinema
5	W5	Rainy	No	Rich	Stay In
6	W6	Rainy	Yes	Poor	Cinema
7	W7	Windy	No	Poor	Cinema
8	W8	Windy	No	Rich	Shopping
9	W9	Windy	Yes	Rich	Cinema
10	W10	Sunny	No	Rich	Tennis

Excluded Feature

Finding GINI Index of the System

Gini(System) = 1 - Gini(Decision=Cinema) - Gini(Decision = Tennis) - Gini (Decision = Stay In) - Gini (Decision = Shoppin)
Gini(System) = 1 - (6/10)^2 - (2/10)^2 - (1/10)^2 - (1/10)^2
= 0.58

Finding GINI index of the Feature = Money

Gini(Feature=Money) = weight average of { Gini(Money,Rich) , Gini(Money,Poor) }

Gini(Money, Rich) = 1 - (3/7)^2 - (2/7)^2 - (1/7)^2 - (1/7)^2
0.6938776

Gini(Money,Rich)

	Weekened	Weather	Parents	Money	Decision
1	W1	Sunny	Yes	Rich	Cinema
2	W2	Sunny	No	Rich	Tennis
3	W3	Windy	Yes	Rich	Cinema
4	W5	Rainy	No	Rich	Stay In
5	W8	Windy	No	Rich	Shopping
6	W9	Windy	Yes	Rich	Cinema
7	W10	Sunny	No	Rich	Tennis

Gini(Money, Poor) = 1 - (3/3)^2
0

	Weekened	Weather	Parents	Money	Decision
1	W4	Rainy	Yes	Poor	Cinema
2	W6	Rainy	Yes	Poor	Cinema
3	W7	Windy	No	Poor	Cinema

Gini(Feature=Money) = Gini(Money,Rich) * Ratio of Rich + Gini(Money,Poor) * Ratio
0.4857143

Finding GINI index of the Feature = Weather

Gini (Feature=Weath)= average weighted[Gini(Weather=Sunny) , Gini(Weather=Windy) , Gini(Weather=Rainy)]

Gini (Weather=Sunny)= 1 - (1/3)^2 - (2/3)^2
= 0.4444444

	Weekened	Weather	Parents	Money	Decision
1	W1	Sunny	Yes	Rich	Cinema
2	W2	Sunny	No	Rich	Tennis
3	W10	Sunny	No	Rich	Tennis

Gini(Weahter=Windy)= 1 - (3/4)^2 - (1/4)^2
= 0.375

	Weekened	Weather	Parents	Money	Decision
1	W3	Windy	Yes	Rich	Cinema
2	W7	Windy	No	Poor	Cinema
4	W8	Windy	No	Rich	Shopping
5	W9	Windy	Yes	Rich	Cinema

Gini(Weahter=rainy)= 1 - (2/3)^2 - (1/3)^2
0.4444444

	Weekened	Weather	Parents	Money	Decision
1	W4	Rainy	Yes	Poor	Cinema
2	W5	Rainy	No	Rich	Stay In
3	W6	Rainy	Yes	Poor	Cinema

Gini(Feature=Weather) = (ratio of sunny in dataset) * Gini(Weather=Sunny) + 0.13333
(ratio of windy in dataset) * Gini(Weather=windy) + 0.15
(ratio of rainy in dataset) * Gini(Weather=Rainy) 0.13333

0.4166667

Finding GINI index of the Feature = Parents

Gini(Feature=parents) = Weight Average [Gini(Parents=Yes) , Gini(Parents=NO)]

Gini(Parents=Yes)= 1 - (5/5)^2
0

	Weekened	Weather	Parents	Money	Decision
1	W1	Sunny	Yes	Rich	Cinema
2	W3	Windy	Yes	Rich	Cinema
3	W4	Rainy	Yes	Poor	Cinema
4	W6	Rainy	Yes	Poor	Cinema
5	W9	Windy	Yes	Rich	Cinema

Gini (parents=No)= 1-(2/5)^2 - (1/5)^2-(1/5)^2-(1/5)^2
0.72

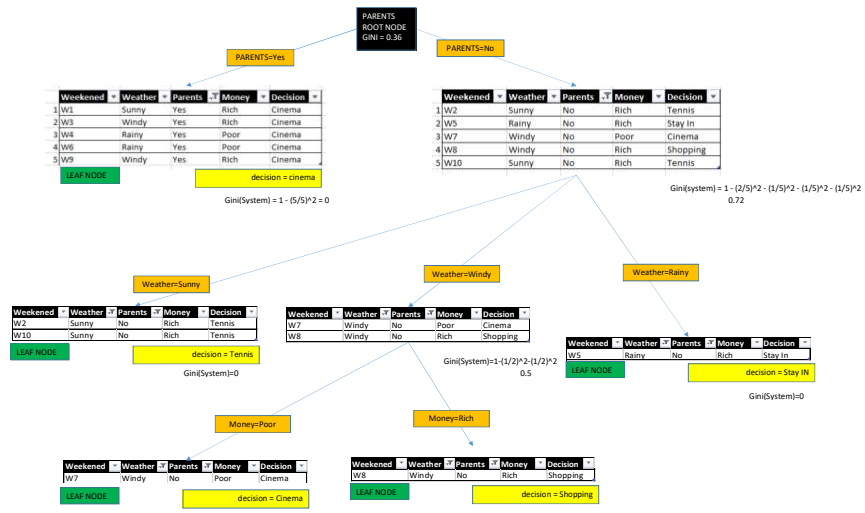
	Weekened	Weather	Parents	Money	Decision
1	W2	Sunny	No	Rich	Tennis
2	W5	Rainy	No	Rich	Stay In
3	W7	Windy	No	Poor	Cinema
4	W8	Windy	No	Rich	Shopping
5	W10	Sunny	No	Rich	Tennis

Gini(Feature=Parents)= ratio of "YES" in dataset * Gini(Parent=yes) + Ration of "NO" in dataset * Gini(Parents=NO)

Gini(Feature=Parents)= 0.36

Assesment of Root Node

	Weather	Parent	Money	Result	Remarks
Gini Index	0.4166667	0.36	0.48571	Parent	Feature with minimum GINI Index has the max info gain



SourcE: (3) Build Decision Tree using Gini Index Solved Numerical Example Machine Learning by Dr. Mahesh Huddar - YouTube