

Decision Tree

Assignment

About the data:

Let's consider a Company dataset with around 10 variables and 400 records.

The attributes are as follows:

- Sales
- Competitor Price
- Income
- Advertising
- Population
- Price
- Shelf Location at stores
- Age
- Education
- Urban
- US

The company dataset looks like this:

Sales	CompPrice	Income	Advertising	Population	Price	ShelveLoc	Age	Education	Urban	US
9.5	138	73	11	276	120	Bad	42	17	Yes	Yes
11.22	111	48	16	260	83	Good	65	10	Yes	Yes
10.06	113	35	10	269	80	Medium	59	12	Yes	Yes
7.4	117	100	4	466	97	Medium	55	14	Yes	Yes
4.15	141	64	3	340	128	Bad	38	13	Yes	No

Problem Statement:

A cloth manufacturing company is interested to know about the segment or attributes causes high sale.

Approach - A decision tree can be built with target variable Sale (we will first convert it in categorical variable) & all other variable will be independent in the analysis.