Categorical Feature Encoding

HOST : Deepesh Mhatre
Github :
github.com/deepeshdm

Categorical Variable

Ordinal
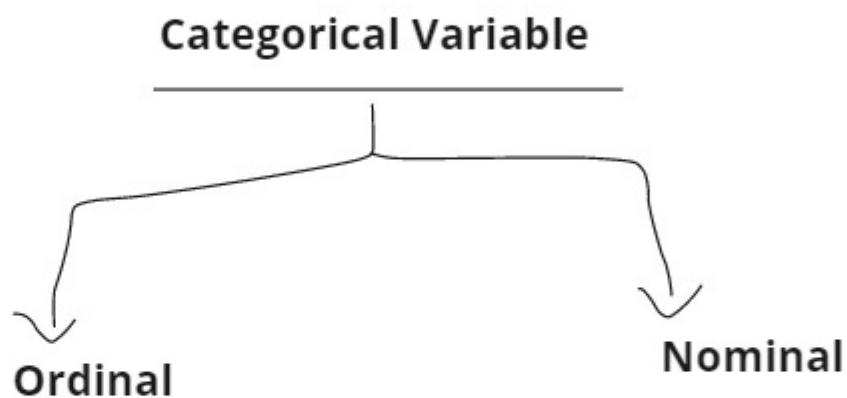
Categorical variables with natural order among values

Eg - Customer Rating can be (good,very good,excellent)

Nominal

Categorical variables with NO natural order among values

Eg - (Cow,Dog,Cat)

(Data of a Hotel Waiter)

Good -- 1
Very Good -- 2
Bad --- 3
Excellent --- 4

Name	Gender	Customer Rating	New column	
John	Male	Good	1	
Ammy	Female	Very Good	2	
Molly	Female	Bad	3	
Paul	Male	Excellent	4	
Batman	Male	Good	1	

Two ways to convert Categorical Features into Numerical

ONE-HOT ENCODING

LABEL ENCODING

NOTE : Usually we use ONE-HOT Encoding for Nominal Features & LABEL Encoding for
Ordinal Features. If the Cardinality of a feature is high,avoid using ONE-HOT
encoding.

DEMO OF FEATURE ENCODING

miro