

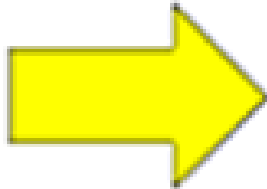


Ordinal encoding

Ordinal encoding: definition

- Ordinal encoding consist in replacing the categories by digits from 1 to n (or 0 to $n-1$, depending the implementation), where n is the number of distinct categories of the variable.
- The numbers are assigned arbitrarily.
- This encoding method allows for quick benchmarking of machine learning models.

Ordinal encoding : example



Color
Red
Red
Yellow
Green
Yellow

Color
1
1
2
3
2



Ordinal encoding: Advantages

- Straightforward to implement
- Does not expand the feature space
- Can work well enough with tree based algorithms



Ordinal encoding: Limitations

- Does not add extra information while encoding
- Not suitable for linear models
- Does not handle new categories in test set automatically

THANK YOU

www.trainindata.com