Train In Data

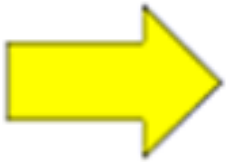
Ordered ordinal encoding

Ordered ordinal encoding: definition

- Categories are replaced by integers from 1 to k, where k is the number of distinct categories in the variable, but this numbering is informed by the mean of the target for each category.

Ordered ordinal encoding: example

Color
Red
Red
Yellow
Green
Yellow



Target	Color
0	2
1	2
1	1
0	3
1	1

Ordered ordinal encoding: Advantages

- Straightforward to implement
- Does not expand the feature space
- Creates monotonic relationship between categories and target

Ordered ordinal encoding: Limitations

- May lead to over-fitting
- Difficult to implement together with cross-validation with current libraries

Train In Data

THANK YOU

www.trainindata.com