How do we know we have the good predicators

We want features that are related to the target but not to each other. How

do we know if features are related?

We can use the correlation variables:

Define X and Y, Two classes.

Case:

X is continuous and Y is a class.

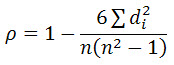
Ordinal data and nominal data

The rank of data: Spearman Coefficient Example

Focus on the variable direction and correlated, not the distribution.

Causation:

Some exceptions.



The ecoefficiency is 6.

**How do we know we’ve done well at classification?**

Build test, see the accuracy.