



Mutual Information



Mutual Information

Mutual information is a measure of the mutual dependence of 2 variables.

In other words, the mutual information quantifies the "amount of information" gained about one random variable through observing the other random variable.

Mutual Information

$$I(X; Y) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p_{(X,Y)}(x, y) \log \left(\frac{p_{(X,Y)}(x, y)}{p_X(x) p_Y(y)} \right)$$

- I is the mutual information
- $p(X,Y)$ is the probability of X and Y occurring together
- $p(X)$ is the probability of X
- $p(Y)$ is the probability of Y

Mutual Information

$$I(X; Y) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p_{(X,Y)}(x, y) \log \left(\frac{p_{(X,Y)}(x, y)}{p_X(x) p_Y(y)} \right)$$

- Determines how similar the joint distribution $p(X,Y)$ is to the products of individual distributions $p(X)p(Y)$
- If X and Y are independent, their I is zero
- If X is deterministic of Y , the I is the uncertainty in X .



Mutual Information: intuition

<https://math.stackexchange.com/questions/2298115/calculate-mutual-information-using-joint-probability-distribution-table?newreg=af0608326cf744c6b843478b84d66587>

Mutual Information: Scikit-learn

Step 1

- Mutual_info_classif
 - https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.mutual_info_classif.html
- Mutual_info_regression
 - https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.mutual_info_regression.htm
- Rank the features

Mutual Information: Scikit-learn

Step 2

- SelectKBest
 - Selects top k highest ranking features
- SelectPercentile
 - Selects features in the top percentile. User defines which percentile, e.g., `top 10th`
- Select the features

Mutual Information: Scikit-learn

Feat 1	Feat 2	Feat 3	Target
F1,1	F2,1	F3,1	0
F1,2	F2,2	F3,2	1
F1,3	F2,3	F3,3	0
F1,4	F2,4	F3,4	0
F1,5	F2,5	F3,5	1
F1,6	F2,6	F3,6	0
F1,7	F2,7	F3,7	1

- **Step 1:** determine mutual information between each feature and the target
 - `Mutual_info_classif`

Mutual Information: Scikit-learn

Feat 1	Feat 2	Feat 3	Target
F1,1	F2,1	F3,1	0
F1,2	F2,2	F3,2	1
F1,3	F2,3	F3,3	0
F1,4	F2,4	F3,4	0
F1,5	F2,5	F3,5	1
F1,6	F2,6	F3,6	0
F1,7	F2,7	F3,7	1

0.31

0.45

0.20

Mutual information

- **Step 1:** determine mutual information between each feature and the target
 - `Mutual_info_classif`

Mutual Information: Scikit-learn

Feat 1	Feat 2	Feat 3	Target
F1,1	F2,1	F3,1	0
F1,2	F2,2	F3,2	1
F1,3	F2,3	F3,3	0
F1,4	F2,4	F3,4	0
F1,5	F2,5	F3,5	1
F1,6	F2,6	F3,6	0
F1,7	F2,7	F3,7	1

0.31

0.45

0.20

Mutual information

- **Step 2:** rank the features based on the mutual information
 - Feat 2
 - Feat 1
 - Feat 3

Mutual Information: Scikit-learn

Feat 1	Feat 2	Feat 3	Target
F1,1	F2,1	F3,1	0
F1,2	F2,2	F3,2	1
F1,3	F2,3	F3,3	0
F1,4	F2,4	F3,4	0
F1,5	F2,5	F3,5	1
F1,6	F2,6	F3,6	0
F1,7	F2,7	F3,7	1

0.31

0.45

0.20

Mutual information

- **Step 3:** select top k ranking feature
 - K=1
 - Feat 2

THANK YOU

www.trainindata.com