# Feature Selection Based on Mutual Information

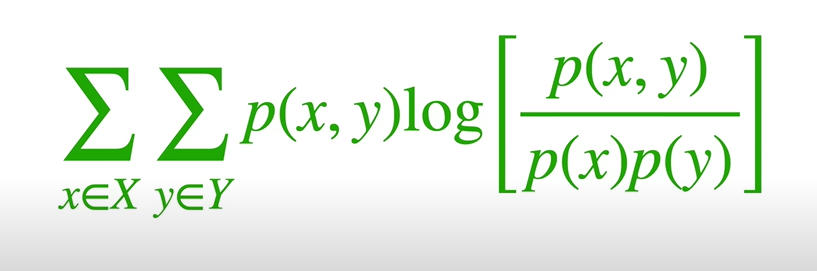
An important part of machine learning process is to understand data and its attributes. There are several attributes present in data, but understanding the important ones is very important. Feature selection is an important process in data cleaning. One such feature selection technique is Mutual Information.

This technique reduces the size of the input feature set. It retains the class discriminatory information, thus ensuring data can be used for classification problems.

Mutual information (MI) between two random variables is a non-negative value, which measures the dependency between the variables. It is equal to zero if and only if two random variables are independent, and higher values mean higher dependency.

In short, it is the amount of information one variable gives about the other.

Let us take two random variables there mutual information between them will be zero if and only if the variables are completely independent otherwise the mutual information between them would be symmetric and non-negative.



<https://guhanesvar.medium.com/feature-selection-based-on-mutual-information-gain-for-classification-and-regression-d0f86ea5262a>

<https://towardsdatascience.com/introduction-to-data-preprocessing-in-machine-learning-a9fa83a5dc9d>

<https://thuijskens.github.io/2017/10/07/feature-selection/>

<https://www.youtube.com/watch?v=eJIp_mgVLwE>

<https://www.youtube.com/watch?v=U9h1xkNELvY>