

STANDARDIZATION

Standardization means to rescale the data to have a mean of 0 and a standard deviation of 1 to fit standard normal distribution.



STANDARD SCALER

ROBUST SCALER

QUANTILE TRANSFORMER

LOG TRANSFORMER

POWER TRANSFORMER SCALER

V/S

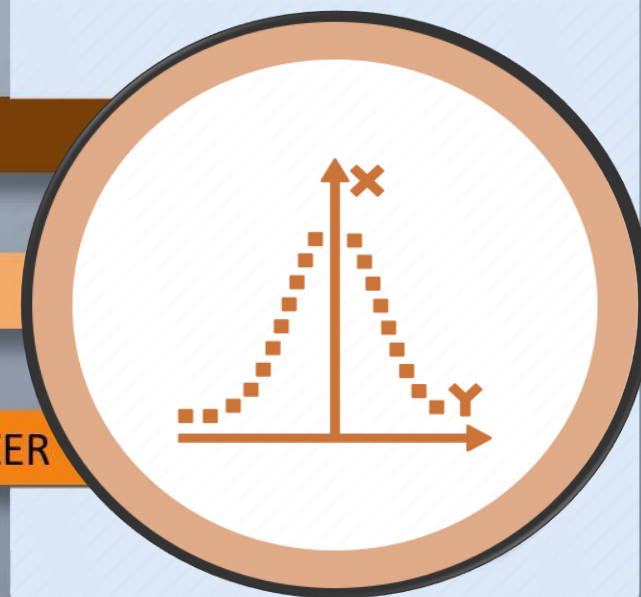
NORMALIZATION

Normalization means to rescale the data into a range of [0,1].

MIN MAX SCALER

MAX ABS SCALER

UNIT VECTOR NORMALIZER



When Should You Use Normalization?

- **Normalization** is a good technique to use when you do not know the distribution of your data or when you know the distribution is not Gaussian (a bell curve).
- **Normalization** is useful when your data has varying scales and the algorithm you are using does not make assumptions about the distribution of your data, such as **k-nearest neighbors** and **artificial neural networks**.

When Should You Use Standardization?

- **Standardization** assumes that your data has a Gaussian (bell curve) distribution. This does not strictly have to be true, but the technique is more effective if your attribute distribution is Gaussian.
- **Standardization** is useful when your data has varying scales and the algorithm you are using does make assumptions about your data having a Gaussian distribution, such as **linear regression**, **logistic regression**, and **linear discriminant analysis**.