Experiment no. : 10

Title :

Perform standardization operation on given dataset.

Theory :

Standardizing a dataset involves rescaling the distribution of values so that mean of observed values is 0 and S.D is 1.

This can be thought of as subtracting mean value or centering data.

Like normalization, standardization can be useful and even required in some M.L algorithm. When your Time series data has input values with different scales.

standardization assumes that your observations fit a Gaussian distribution with a well behaved mean and S.D you can still standardize your time series data if this expectations is not met, but you may not get reliable results.

This includes algorithm like support vector machines linear and logistic regressions and other algorithm that assume or have improved performance with Gaussian data.

standardization requires that you know or are able to accurately estimate mean and S.D of observable values. you may be able to estimate these values from your training data

A value is standardized as follows

$$Y = (X - \text{mean}) / \text{standard\_deviation}$$

where mean is calculated as

Mean = sum(x) / count(x)

and standard_deviation is calculated as :

standard_deviation = sqrt(sum((x - Mean)^2) / count(x))

conclusion :-

In this experiment we have studied that how to apply standardization operation on given dataset.