Feature Scaling

When to apply feature scaling?

1.scaling will be effective when you have variations within a data.

2. when the algorithm is based on distance formula.

Feature scaling is an important preprocessing step for machinelearning models, in which we transform the data into a better version and scale the dataset into a finite range. Normalization or Standardization are the two main techniques that #datascientist use.

Normalization is the process of rescaling values into a common scale between a range of [0,1] or [-1,1] by using the MinMaxScaler transformer from the Scikit-Learn package.

Standardization, also known as Z-Score Normalization, involves rescaling the data based on the standard normal distribution and is performed using the StandardScaler transformer in Scikit-Learn

Normalization is useful when data has varying scales and you don't know the distribution of your data and is useful when the algorithm you are using does not make assumptions about the distribution of your data, such as k-nearest neighbors and artificial neural networks.

Standardization is useful when data has varying scales and

when you have assumed your data has a Gaussian distribution

and is useful when the algorithm you are using does make assumptions about the distribution, such as linear regression, logistic regression, and linear discriminant analysis.