-> Scaling :-

The technique to standardize the independent features present in the data in a "fixed range". (0 - 1 or -1 to +1)

In ML algo tend to give more weight to values with higher magnitude irrespective of the unit, so its better to standarize values before applying algorithm.

Date: __/__/____

↳ Feature scaling is part of data preprocessing.

↳ Only done after splitting data on train data as we want test data to be untouched.

↳ It significantly increses the performance of algo as now values are closely located (0 to 1)

Types of scaling :

1) Normalization (Min-Max) [0-1] $X' = \frac{X - X_{min}}{X_{max} - X_{min}}$

2) Standarization (centered around mean with unit SD(σ)) $X' = \frac{X - \mu}{\sigma}$

↳ Scaling should be done in algo which are distance based like KKN, K-means or SVM, where it does not effect tree based algos.

↳ Having feature on similar scale can help the gradient descent converge more quickly towards minima.