EXPT. NO. | NAME | Page No. | Date | YOUVA

M T W T F S S

* Feature Scaling :- (Standardization) (Day: 24)

"Feature Scaling is a technique to standardize the independent features present in the data in a fixed range."

→ Feature Scaling is the process of transforming numerical features into a common scale without distorting differences in the range of values.

→ For eg:- If one feature is Salary = 50,000 and another is Age = 50, the large scale of salary may overshadow age in certain algo. By ~~so~~ scaling, both features are brought to a comparable range (eg: 0-1 or with mean=0 and SD = 1).

General FS

→ If apde aa vastu solve na karye, then amuk apda ML algo. ma train khotu thase and it will not give you the accurate results for this

⇒ Types of Feature Scaling :-
(1) Standardisation.
(2) Normalisation.

⇒ Standardisation :- (also known as Z-score Normalization)

→ To do this the formula for x column is :-

$$x'_i = \frac{x_i - \bar{x}}{\sigma}$$

$\bar{x}$ → mean.

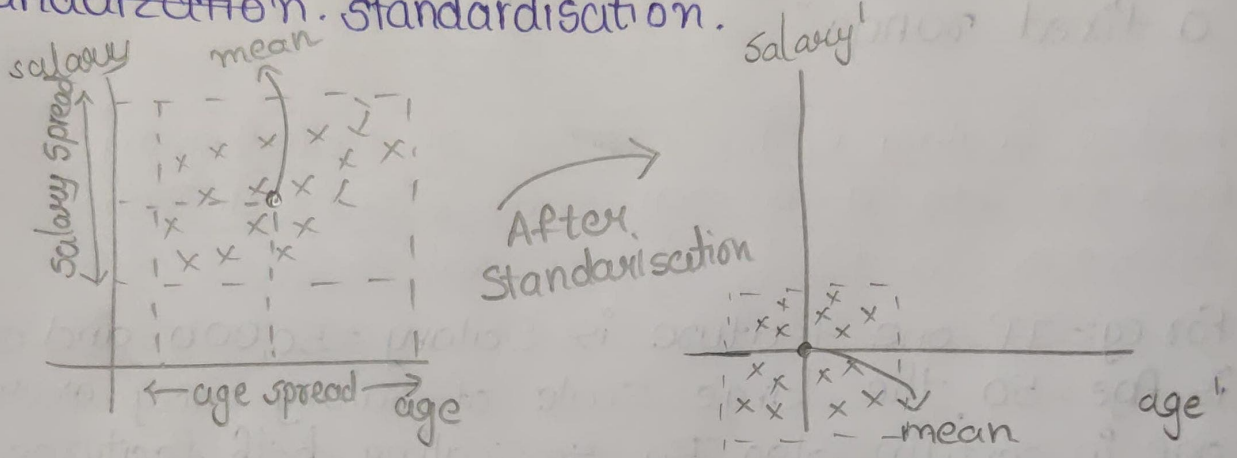
σ → Standard deviation

Teacher's Signature:

→ Now the mean of the new column created will be 0 and standard deviation will be 1

⇒ mean = 0, ~~sta~~ SD = 1.
(μ) (σ)

This characteristics is fixed when you apply ~~standaizeation~~. standardisation.



σ = 1. (mean centring)

Both the salary and age spread should be within 1 so σ = 1

⇒ What if I have outlier in my data?

→ if apde outliers wala column ma standardisation lagavye che to eno impact ocho nathi thatu, e outlier ni jem j behave kare che.

→ we have to ~~he~~ handle the outlier explicitly.

⇒ When to ~~y~~ use?

→ we can use it any algo., there is no disadvantage of using it. but it is more advantage in ~~st~~ certain algo. such as logisticRegression, Your accuracy will increase after scaling.

→ Use you using any of the below algo. use standardisation definitely.

(1) K-Means (3) PCA ~~(4)~~ (5) Gradient Descent.
(2) KNN (4) Artifical Neural Networks

EXPT. NO.

NAME

M T W T F S S

Page No.

YOUVA

Date: 12/2/25

→ when you want to find the distance between two points in any of the algo then you must use standardisation.