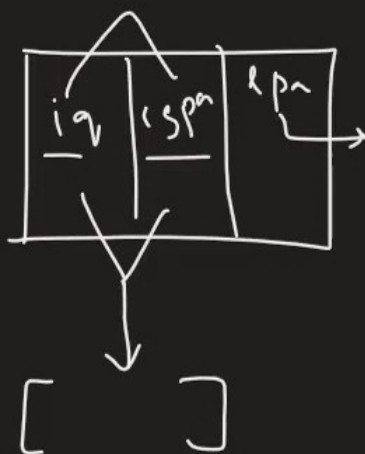




What is Feature Scaling?

Friday, April 9, 2021 4:21 PM

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



100 10

100 Days of ML

Day 6 - Instance Vs...

Day 7 - Challenges...

Day 8 - Application...

Day 9 - MLDLC

Day 10 - Job Roles

Day 11 - Tensors

Day 12 - Setting u...

Day 13 - End to En...

Day 14 - Framing t...

Day 15 - Working...

Day 16 - Working...

Day 17 - API to Pa...

Day 18 - Web Scra...

Day 19 - Understa...

Day 20 - Univariate...

Day 21 - Bivariate...

Day 22 - Pandas Pr...

Day 23 - Feature E...

Day 24 - Standardi...

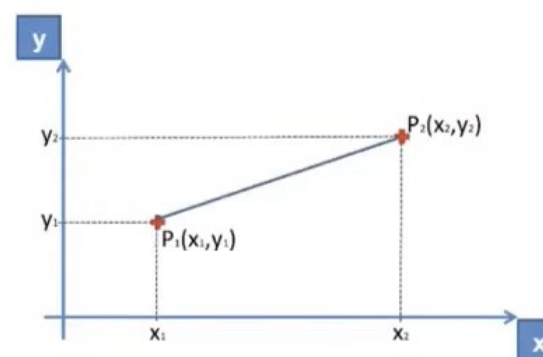
+ Add section

+ Add page



Why do we need Feature Scaling?

Friday, April 9, 2021 4:27 PM



$$\text{Euclidean Distance between } P_1 \text{ and } P_2 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Let x be the no. of Salary and y be the no. of Age

Example: x_1 & y_1 are in row 2, x_2 & y_2 are in row 9

$$(x_2 - x_1)^2 = (83000 - 48000)^2$$

$$= 1225000000$$

$$(y_2 - y_1)^2 = (50 - 27)^2$$

$$= 529$$

age | salary | purchase | Knn

→ 50	83000	1
→ 27	48000	0

+



100 Days of ML

- Day 6 - Instance Vs...
- Day 7 - Challenges...
- Day 8 - Application...
- Day 9 - MLDLC
- Day 10 - Job Roles
- Day 11 - Tensors
- Day 12 - Setting u...
- Day 13 - End to En...
- Day 14 - Framing t...
- Day 15 - Working...
- Day 16 - Working...
- Day 17 - API to Pa...
- Day 18 - Web Scra...
- Day 19 - Understa...
- Day 20 - Univariate...
- Day 21 - Bivariate...
- Day 22 - Pandas Pr...
- Day 23 - Feature E...
- Day 24 - Standardi...

What is Feature Scaling?

Why do we need Feature Scaling?

Types of Feature Scaling

Standardization - Intuition

Example

Impact of Outliers

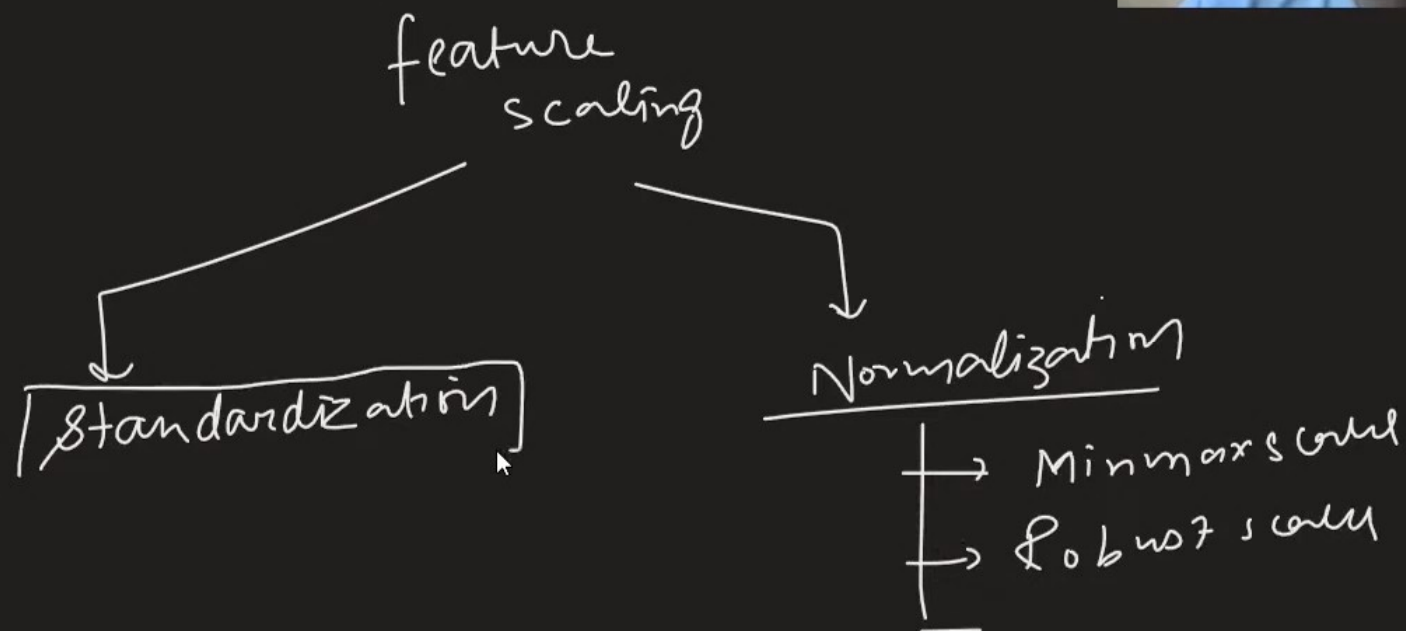
When to use Standardization?

+ Add section

+ Add page

Types of Feature Scaling

Friday, April 9, 2021 4:27 PM





Standardization - Intuition

Friday, April 9, 2021 4:27 PM

Also called as Z-score Normalization

X_i Age salary

27	x_1	2.3
15	x_2	-1.2
33	x_1	-
63	-	-
90	-	-
05	-	-
:	-	-
500 values		

32-

$$\frac{15 - 32}{10}$$

age'

$$X_i' = \frac{X_i - \text{mean}}{\sigma}$$

mean

σ SD

$$X_1' = \frac{27 - 32}{10}$$

100 Days of ML

Day 6 - Instance Vs...

Day 7 - Challenges...

Day 8 - Application...

Day 9 - MLDLC

Day 10 - Job Roles

Day 11 - Tensors

Day 12 - Setting u...

Day 13 - End to En...

Day 14 - Framing t...

Day 15 - Working...

Day 16 - Working...

Day 17 - API to Pa...

Day 18 - Web Scra...

Day 19 - Understa...

Day 20 - Univariate...

Day 21 - Bivariate...

Day 22 - Pandas Pr...

Day 23 - Feature E...

Day 24 - Standardi...

+ Add section

+ Add page



100 Days of ML

- Day 6 - Instance Vs...
- Day 7 - Challenges...
- Day 8 - Application...
- Day 9 - MLDLC
- Day 10 - Job Roles
- Day 11 - Tensors
- Day 12 - Setting u...
- Day 13 - End to En...
- Day 14 - Framing t...
- Day 15 - Working...
- Day 16 - Working...
- Day 17 - API to Pa...
- Day 18 - Web Scra...
- Day 19 - Understa...
- Day 20 - Univariate...
- Day 21 - Bivariate...
- Day 22 - Pandas Pr...
- Day 23 - Feature E...
- Day 24 - Standardi...

+ Add section + Add page

When to use Standardization?

Friday, April 9, 2021 4:28 PM

Algorithm(s)	Reason of applying feature scaling
1. <u>K-Means</u>	Use the Euclidean distance measure.
2. <u>K-Nearest-Neighbours</u>	Measure the distances between pairs of samples and these distances are influenced by the measurement units
3. <u>Principal Component Analysis (PCA)</u>	Try to get the feature with maximum variance
4. Artificial Neural Network	Apply Gradient Descent
5. <u>Gradient Descent</u>	Theta calculation becomes faster after feature scaling and the learning rate in the update equation of Stochastic Gradient Descent is the same for every parameter

