



Normalization

Saturday, April 10, 2021 1:15 PM

Normalization is a technique often applied as part of data preparation for machine learning. The goal of normalization is to change the values of numeric columns in the dataset to use a common scale, without distorting differences in the ranges of values or losing information

+ MinMaxScaling
+ Mean normalization
+ Max absolute
- Robust scaling

Normalization

MinMaxScaling - Intuition

Code Example

Mean Normalization

MaxAbsScaling

Robust Scaling

Normalization Vs Standardization

100 Days of ML

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...

Day 25 - Normaliza...

+ Add section

+ Add page



MinMaxScaling - Intuition

Saturday, April 10, 2021 1:16 PM

weight

130

67

81

61

32

54

...

...

100 wt

Normalize

Min max scaling

min = 32

max = 130

67 - 32

130 - 32

$x_{min} = 0$

$[0, 1]$

$$x_i' = \frac{x_i - x_{min}}{x_{max} - x_{min}} = \frac{x_i - 0}{130 - 32}$$

$$= \frac{130 - 32}{130 - 32} = 1$$

- 100 Days of ML
- 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...
- Day 25 - Normaliza...

+ Add section

+ Add page



Home

Insert

Draw

View

Help



Shapes

Link to Shape

Link to Text



100 Days of ML

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...

Day 25 - Normaliza...

Normalization

MinMaxScaling - Intuition

Code Example

Mean Normalization

MaxAbsScaling

Robust Scaling

Normalization Vs Standardization

+ Add section

+ Add page

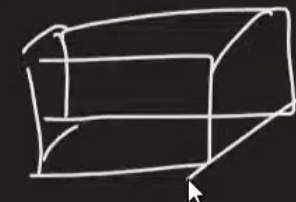
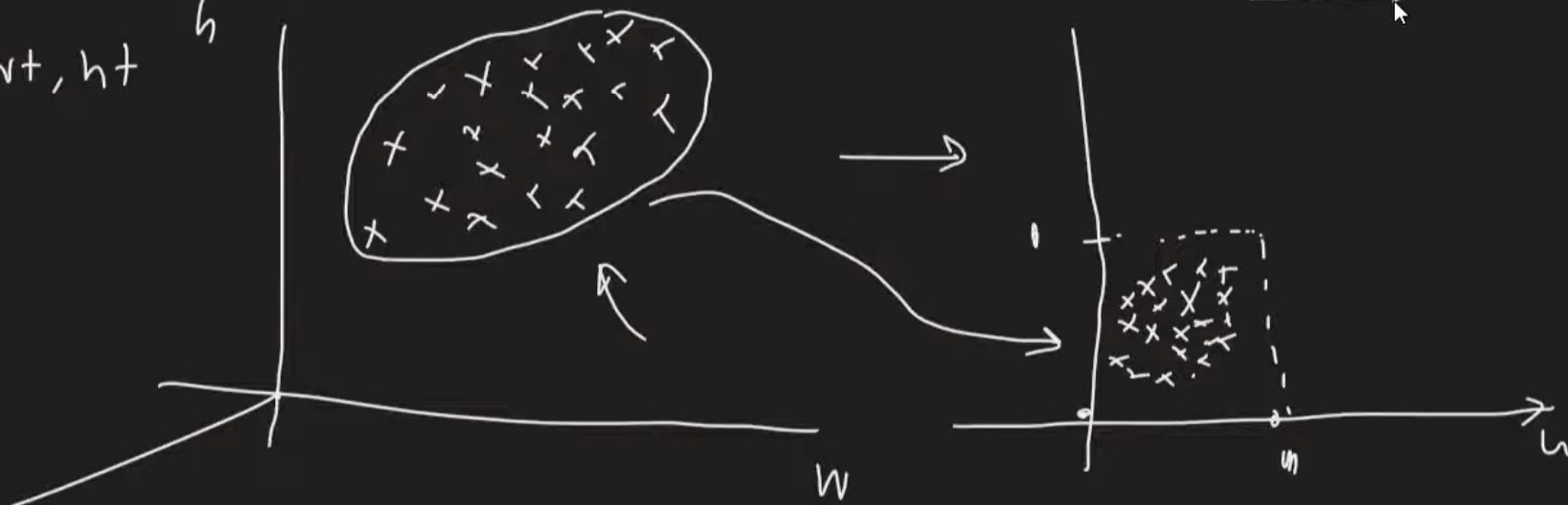
32
54
... $\min = 32$ $\max = 130$ $67 - 32$ $130 - 32$ X_{max}

$$= \frac{130 - 32}{130 - 32} = 1$$

(100)
wt

wt, ht

h





Mean Normalization

Saturday, April 10, 2021 1:16 PM

wt
200
100

normal

$\text{val} < \text{mean}$

$-\text{val} \mid \text{Lve}$

$$X'_i = \frac{X_i - X_{\text{mean}}}{X_{\text{max}} - X_{\text{min}}} [-1 \text{ to } 1]$$

mean center

sklearn →

centered data

Standardization



100 Days of ML

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...

Day 25 - Normaliza...

+ Add section

+ Add page



100 Days of ML

- 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...
- Day 25 - Normaliza...

Normalization

MinMaxScaling - Intuition

Code Example

Mean Normalization

MaxAbsScaling

Robust Scaling

Normalization Vs Standardization

MaxAbsScaling

Saturday, April 10, 2021

1:17 PM

Wt
200
100
300



$$X'_i = \frac{X_i}{|X_{max}|}$$



MaxAbsScaler

+ sparse data → (0s)

+ Add section

+ Add page



Robust Scaling

Saturday, April 10, 2021 1:17 PM

wt

200

300

100

$$X'_i = \frac{X_i - X_{\text{median}}}{\text{IQR} \{ 75^{\text{th}} \text{ per} - 25^{\text{th}} \text{ per} \}}$$

Robust scale

→ Robust to outliers





Normalization Vs Standardization

Saturday, April 10, 2021 1:17 PM

1) Is feature scaling required?

2) MinMax → image → CNN → 0-255 min max

