Normalization

1. Normalization

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.

- Eliminate unit just keep magnitude.

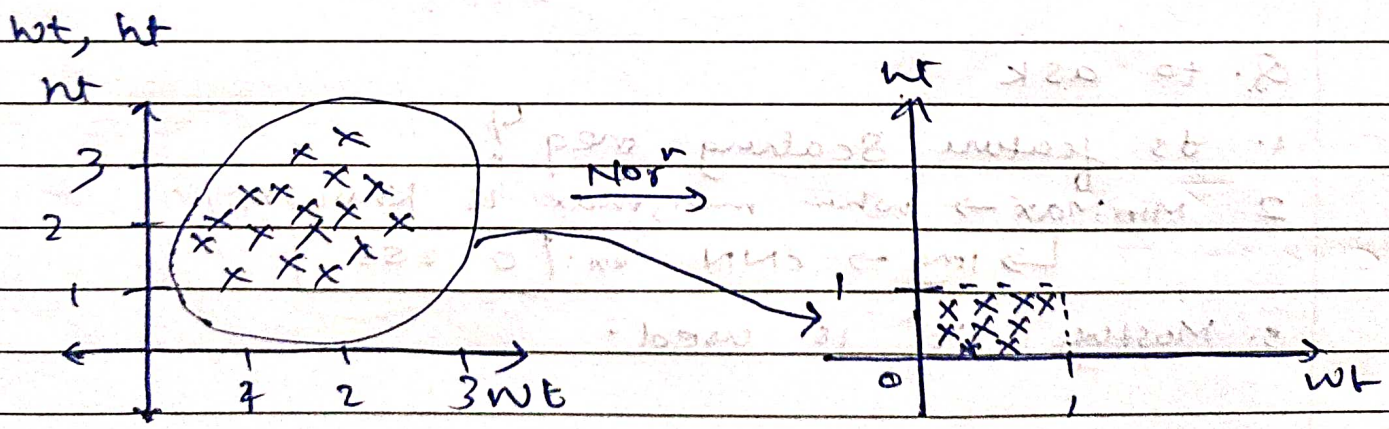
2. MinMax Scaling

Weight
130
67
60
90
⋮
100 values

weight aft. Norm^{n}

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

range [0 - 1]



3. Mean Normalization.

$$x_i' = \frac{x_i - x_{mean}}{x_{max} - x_{min}} \quad [-1 \text{ to } 1]$$

Usefull → when centered data req.

4. MaxAbsScaler

$$x_i' = \frac{x_i}{|x_{max}|}$$

usefull → when sparse data [0's present in large]

5. Robust Scaling

$$x_i' = \frac{x_i - x_{median}}{IQR \{75^{th} \%. - 25^{th} \%.\}}$$

Usefull → When there is lot of Outliers

- Norm^n vs Stand^n

Q. to ask

1. Is feature Scaling req?
2. MinMax → when min, max is known
   ↳ img → CNN ex: [0 - 255]
3. Mostly St^n is used.