Machine Learning Course | Arabic

Data Preprocessing

Level - 01

Link to Lecture on Youtube

04 Standard Deviation, Variance

Variance

The **expected deviation** between values in a given data set. **Measures the spread** of each number from the mean value of the data set.

$$s^2 = \frac{\Sigma (x - \bar{x})^2}{n - 1}$$

Standard Deviation

Gives an indication of **the actual far of a group of samples from their mean**. and is the square root of the variance.

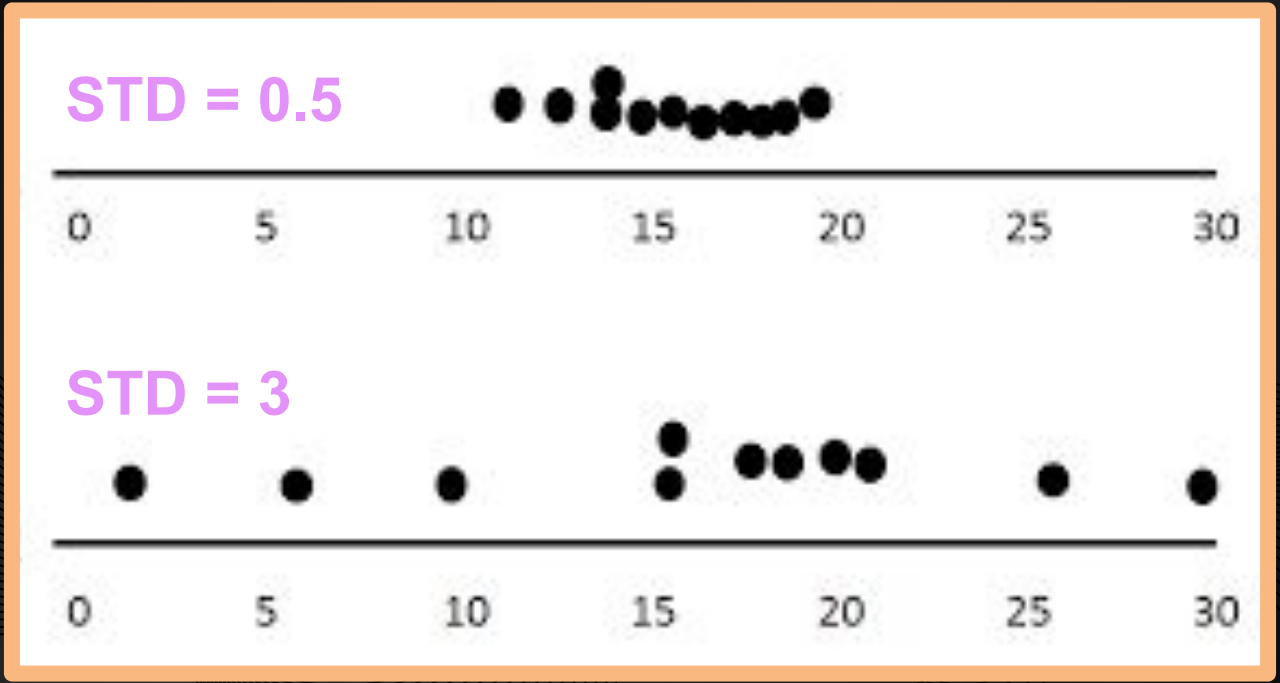
$$s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \overline{x})^2}{N-1}}$$

$\overline{X}$: Mean

N : Number of samples

X : Sample (i)

Different STD values



Variance Vs STD

- The **standard deviation** shows the position of each value from the mean.
- However, a variance is indicated in larger units such as meters squared while the standard deviation is expressed in original units such as meters.
- Variance is used to "**check outliers**" by squaring each value.

  100 ^ 2 = 10000 not like 100 !!!

Ex.

An investor wants to calculate the standard deviation experience with their investment portfolio in the past four months. Here are some historical return numbers:

Month	Return
May-18	15%
Jun-18	-9%
Jul-18	10%
Aug-18	6%

Solution:

1- calculate the mean:

$$\frac{(0.15 - 0.09 + 0.10 + 0.06)}{4} = 0.055$$

So, The arithmetic mean of returns is **5.5%**.

Ex.

An investor wants to calculate the standard deviation experience with their investment portfolio in the past four months. Here are some historical return numbers:

Month	Return
May-18	15%
Jun-18	-9%
Jul-18	10%
Aug-18	6%

Solution:

2- calculate the STD:

$$SD = \sqrt{\frac{(0.15-0.055)^2 + (-0.09-0.055)^2 + (0.10-0.055)^2 + (0.06-0.055)^2}{3}} = 0.1034$$

The standard deviation of returns is **10.34%**.

Thus, the investor now knows that his returns change by about “10%” per month.

The information can be used to adjust and improve the investor's attitude to risk.

Thank You!

Do you have any questions?

Write them in the comments

hozaifazaki99@gmail.com