25/11/25

* A model is considered as overfitting when it's train accuracy is very more than test accuracy.

Lasso:

default $\alpha = 1.0$

```
l = Lasso()
l.fit(x_train, y_train)
y_train_pred = l.predict(x_train)
y_test_pred = l.predict(x_test)
```

It is good to choose small α value.

* When we have to find which α value is better, in that time we use GridSearchCV & RandomSearch CV.
  - ↳ Techniques to find out best hyper parameter value.

Ridge:

default $\alpha = 1.0$

```
r = Ridge()
r.fit(x_train, y_train)
y_train_pred = r.predict(x_train)
y_test_pred = r.predict(x_test)
```

★

Scaling: When multiple features have diff. ranges the ML model decides its importance to avoid it we use scaling.

Types of scaling in machine learning:

* Scaling is applicable only for features, not for target.

f1	f2	f3	f4	f5
0.5	25000	127	3875	0.01
0.3	38000	138	1234	0.003
0.8	19000	149	1001	0.024
0.1	45000	180	1213	0.013
0.9	10000	162	9999	0.005

ML thinks f2, f4, f3, f1, f5
↳ X

Every feature should be equally important.

* In order to give equal importance/priority for each & every feature we go with scaling.

* We can passed scaled values only for equation based algorithms.

Types:

1. Min-Max Scaling (Normalisation)

* It converts values to a fixed range usually 0 to 1.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

f3	Scaled f3.
127	0.3
138	0.5
149	0.7
160	1
112.	0

$X_{min} = 112$

$X_{max} = 160$

2. Standard Scalar (z-score scaling)

It converts data to mean = 0 & SD = 1.

$$X' = \frac{X - \mu}{\sigma}$$

$\mu = 137.2$

$\sigma = 16.7$.

f3	scaled f3	
127	-0.5	-0.6
138	0.04	0.04
149	0.6	0.7
160	1.21	1.36
112	-1.3	-1.5

-0.5454
0.0427
-0.5

$\mu = 0$

$\sigma = 0.99 \approx 1$.

3. Robust Scalar (only when outlier is present)

It uses median & IQR (Interquartile Range)

No fixed range.

$$X' = \frac{X - Median}{IQR}$$

median : 149 | 138

IQR : 35

f3	scaled f3
127	-0.31
138	0
149	0.31
160	0.62
112	-0.74

4. Max - Absolute Scaling (when +ve & -ve data)

* Scales data by dividing by the maximum absolute value.

* keeps sign of data intact.

ranges: -1 to +1

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

5. Log Transformation (in presence of skewness)

* Applies logarithmic transformation to reduce skewness

no fixed range.

$$x' = \log(x+1)$$

* Partially works with outliers.

F3	Max-absdute	Log transformation.
127	0.79	2.10
138	0.86	2.14
149	0.93	2.17
160	1	2.20
112	0.7.	2.05