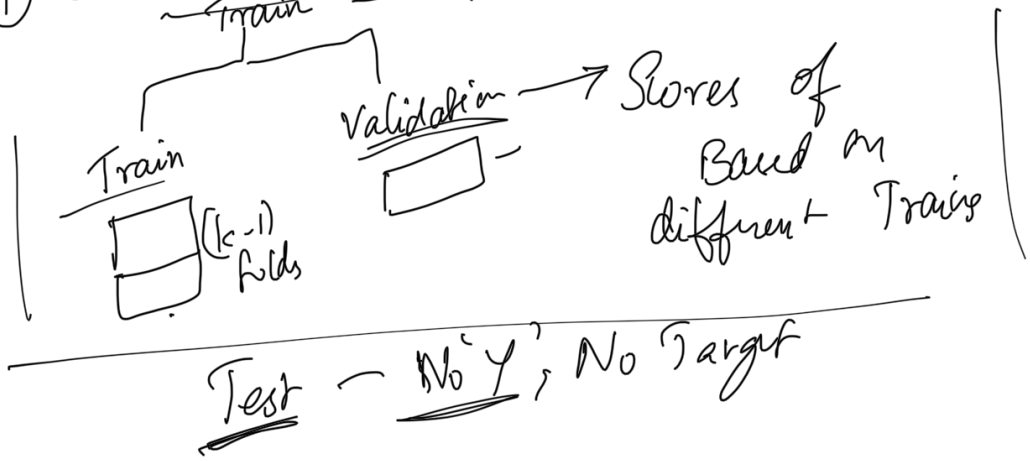
Revisiting 17-12 Bias/variance , Regularization

① Revisit → Crystal Clear. !!

② look at practicals

③ You will be participating 17-18

① Cross Validation → k-folds



Test – No 'Y', No Target

② Bias & Variance

* Bias :- Are the simplifying assumption made by model to make the target easier to learn.

→ Less flexible

→ linear Algorithms — high

↳ ADV:

① Easier to understand

② faster to learn

↳ DIS ADV.

In General ① lower prediction performance on complex problems

Eg) High bias
→ Logistic Reg., LDA, linear Reg., linear Reg. Regression

Eg) Low Bias
→ Decision Tree, KNN, SVM, Deep learning

* Variance:
⇒ The amount that the estimate of Target fn will change if the train data was used.
↳ highly Sensitive

Ideal Case of Variance:-
The model should not try to adapt for each & every training sample. Instead it should find Underlying Pattern !!!

Eg): high variance model:-
In General, non-linear model have high flexibility → high variance.

Bias & Variance

(1) Bias ∝ Variance

(2) Bias ∝ $\frac{1}{\text{Variance}}$

Reg of the Data

Redmin Note 10 Pro

→ Simple Mobile

₹ 8,000

Apple → Iphone 13

₹ 80,000

Overfitting / Under

Sample

3

30,000

Trade-off

(1) High Bias / low Variance.

→ Underfitting

not at all flexible

Train - Error ↑

Test - Error ↑

Variance ↑

(1) complex model

(2) Add more data pts

(3) feature Engineering

② ↑ High Variance / Lower Bias

+ Bias → 1

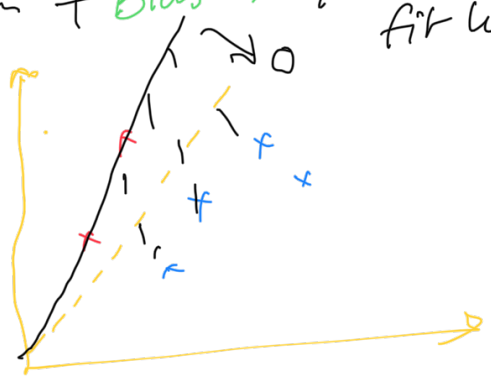
Overfitting

① Train ↑ — Very Good

② Test Not good perfm.

Regularization + Bias → Generalize the fit line

By definition satisfying variance.



— Test

— Train

Regularisation !!

① Using do you improve your Errors with Regularization??

Train Error ⟷ Test Error.

$SS + \lambda(|\text{Co-eff}|)$

$+ \lambda(|\text{Co-eff}|)^2$

Increase λ

Ridge $\lambda \simeq 0$

Variance:

amount of the estimate that the target

λ = 10
λ = 30
so
λ → ∞

Ridge – L2	Lasso – L1	Elastic net
$SS + \lambda(coeff)^2$	$SS + \lambda(\lvert coeff \rvert)$	$\lambda_1(Ridge) + \lambda_2(lasso)$
→ Coeff ≃ 0	→ Coeff → 0	HYBRID.
All the variables will be retained	insignificant var. will be eliminated coeff = 0	If you are not sure use Elastic net

0.00000 ...

$y = c + (coeff_1)x_1 + (coeff_2)x_2 + (coeff_3)x_3 \rightarrow$

Revised Write up
Medium, Analytics Vidh,
Kaggle, Google Blogs,
→ Whatsapp group

k-fold, Leave one out

Break
10:00 am IST