Overfitting
&
Underfitting

Training dataset

High

0.88

x_train, y_train

Testing dataset

low

0.75

x_test, y_test

Overfitting :-

Training data Accuracy high 0.88

Testing data Accuracy low 0.75

0.13

0.90 0.95

0.50 0.61

Underfitting :-

Training dataset Accuracy :- low 0.60%

Testing dataset Accuracy :- low 60%

	V1	V2	V3	T
A		39	78	25
B		38	67	24
C		25	66	23

90 55 60

OF UP UF

T 35 72 23

60 60 57

Test 82 } 11 89 85 } 4

Train

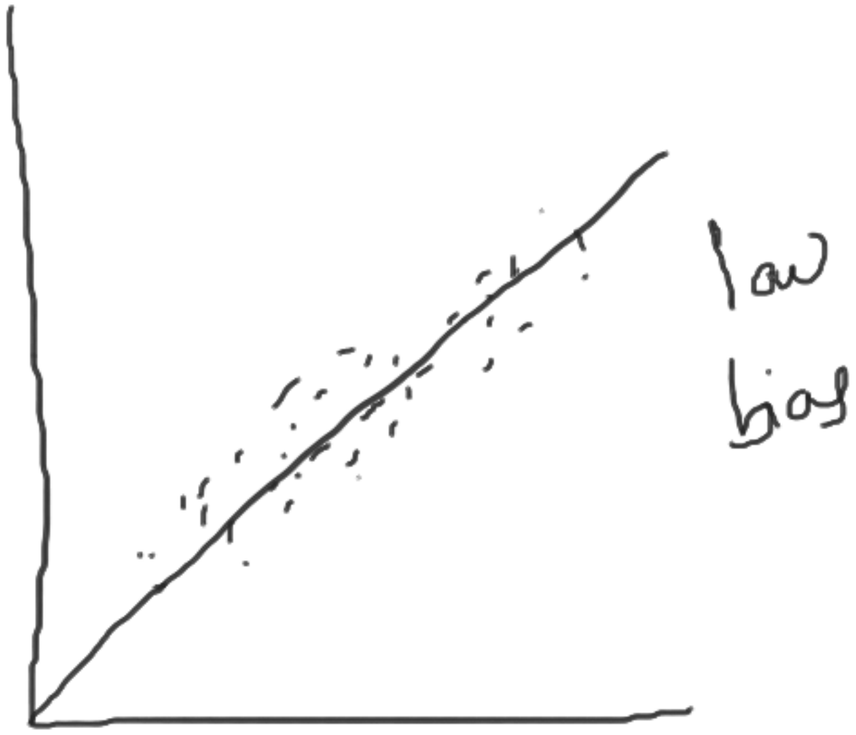
Test 92

Train 91

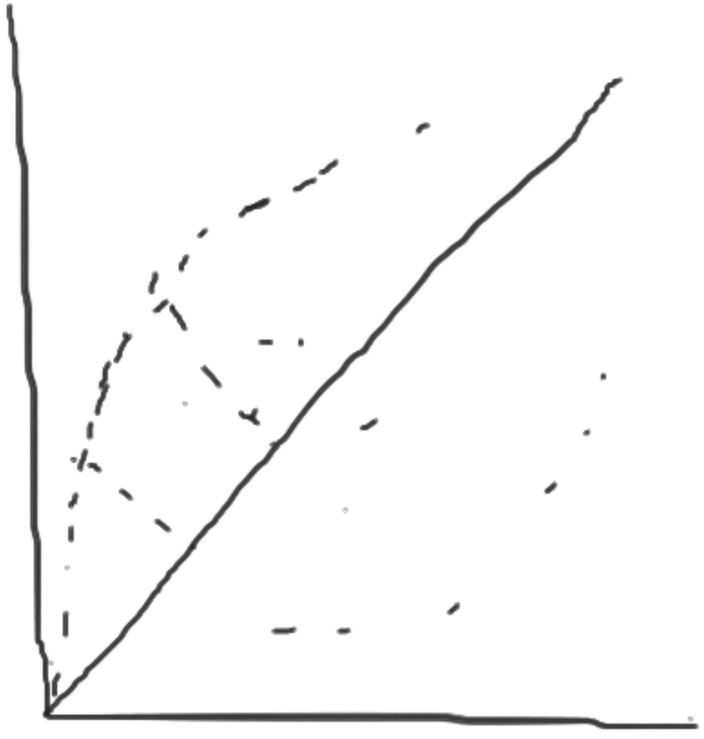
Test 40

Train 30

1) Bias - diff between act & pred

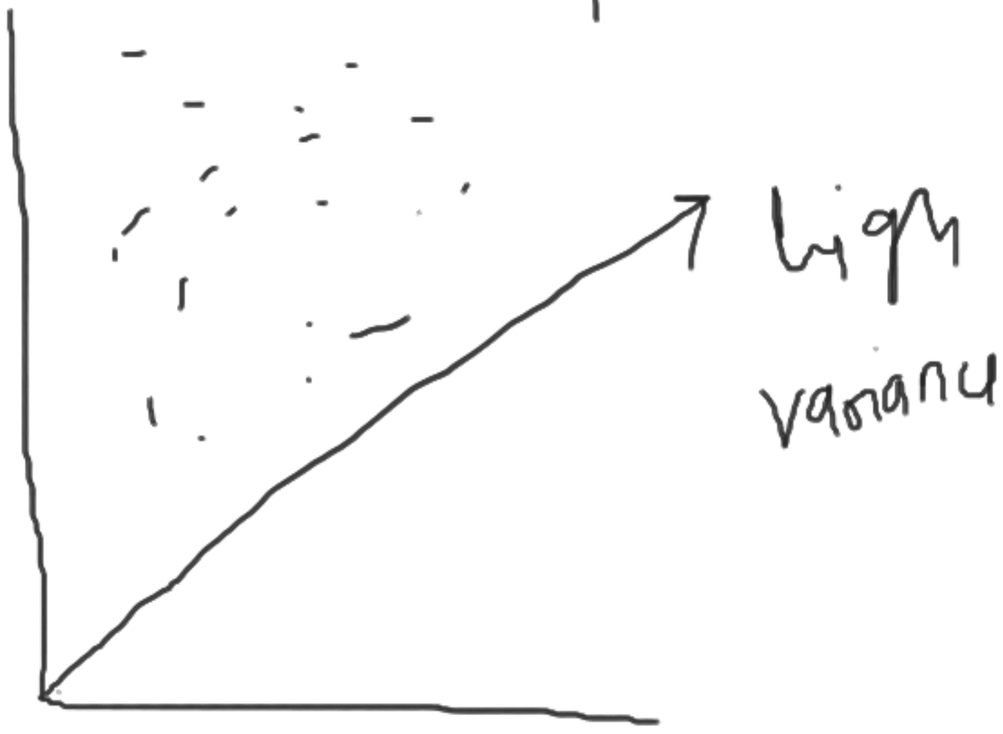


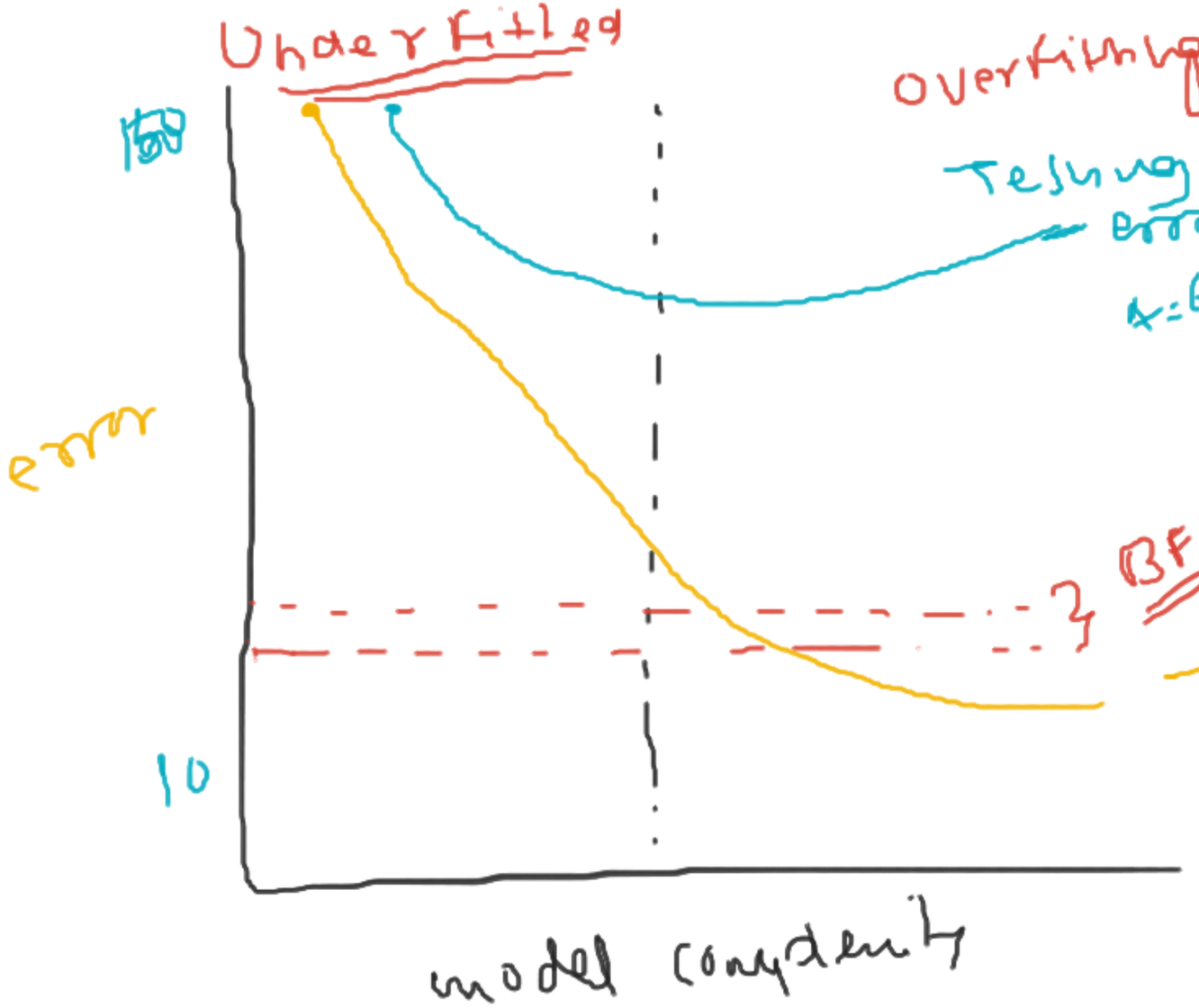
low bias



high bias

2) Variance → How much scattering our values are from each other w.r.t prediction





Acc = 96

mse = 50 %

MA = 50 %

low variance

high variance

High bias

under fitting

low bias

Best model

worst model

over fitting

1st => Technical → project → 3/4 algo

2-4 yrs

→ ML LR
Logi

DC }
RF }

2nd → Managerial → F.S = 60 % F.S = 60 %
Dar = 60 % A.a = 60 %

→ 100 %

9 → 5 4 2n+4 = 28

28 27

M = year

month 2, 3

P.L : 10
S.L

Sat Sunday
29 30 31

3 to 5

16

Company policy

Spine hrm

Cloud hrm Zoho

Absence people

Salary slip

emp UAN

~~darwinbox~~

ID-card

chip

regularize / at update

10-7

9-6

11-8

10-7

12-9

11-10

hybrid

9

online

login
check in
} regularize request, manager
check out
log out

7-5

volant

30

1 - 30

18 - 15

16-Dec to 15-Jan

30 Jan

Product base :- →

Service base :- →

Data leakage

Train

Test

2:

split-

160

40

Train	Test	prod.
80	78	70