$p\uparrow$

$R^2\uparrow$

$$R^2_{adj} = 1 - \frac{(1-R^2)(n-1)}{n-p-1}$$

penalized

minmax

too many feature

complex

where :

$R^2 = R - squared$

n = number of samples/rows in the data set

money heist

p = number of predictors/features

season!

$R^2_{adj} \Downarrow$

$R^2 \sim \Uparrow$

$R^2 \rightarrow \boxed{5 + \text{professor}} + 2$

police

$2 \rightarrow$ imp

Not imp

$\uparrow R^2$

y

mean

nobrainer

SS

↑

$$R^2 = 1 - \frac{RSS}{TSS}$$

RSS — model

TSS → avg

	y	model	nB
1	0.1	0.1	0.25
2	0.2	0.2	0.25
3	0.3	0.3	0.25
4	0.4	0.4	0.25

$$RSS = \sum (y - \hat{y})^2$$

$$\frac{1}{4} = 0.25$$

var

$$TSS = RSS + ESS$$

$$\sum (y - \beta_0 - \beta_1 x_1)^2$$

$$\sum (y - \beta_0 - \beta_1 x_1 - \beta_2 x_2)$$

null hypo → LR → $m = 0$ β_i

$m \neq 0$

R^2 R^2_{adj} features

junk feature $R^2 \uparrow \sim$ $R_{adj} \downarrow$

food

$\uparrow$ immunity = green veg + junk $\Downarrow$

$\Downarrow$ $\Uparrow$

RMSE => ?

avg

whole y'

test

train

$RMSE_{test}$

avg_{tr}

blunder

avg_{test}

apple to apple

evalution metric

$Model_1$ $Model_2$

R^2 MAPE

Interview

Hyperparameters ⟷ **Parameters**

model → fit → data

model data → fit

param → fit

$$CF = \sum (y - \hat{y})^2 + \lambda \beta_i^2$$

1000

10

1

grid search

optimal

max iter = 3000

~~Value~~

LR → $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

β_i

CF → ~~optim~~

min error

LR

lasso &

→ class

lasso = Lasso(alpha=0.1, max_iter= 3000)

var

__init__

value? 0

gridsearch

(Mod)

~~Lasso~~

Penalizing → ~~reducing no. variables~~

Regular

unnecessary

(feature selection)

reducing

$R^2 = 0.9$ adj $R^2 = 0.45$

LASSO

Lasso

cost

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}(y-\hat{y})^2 + r\lambda\sum_{j=1}^{n}|w_j| + \frac{1-r}{2}\lambda\sum_{j=1}^{n} w_j^2$$ → Ridge

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}[y-(wx_i+b)]^2 + r\lambda\sum_{j=1}^{n}|w_j| + \frac{1-r}{2}\lambda\sum_{j=1}^{n} w_j^2$$

r → L1 ratio

min → Lasso / Ridge ← Elasticnet

r = 1 r = 0.5 $\frac{1-0.5}{2} = 0.25$

Hyperparameter Tuning

Too CI

"λ" → confidence → penalizing error

λ → [100 | 10 | 1 | 0.1 | 0.001 | 0.001 | 0.0001]

$c^2 \cdot (y - \hat{y})^2$, "λ"

B^2

model = "EM" → ↓

grid search → Tuning Hyper parameter

fit model → best parameter

energy

10^{-15} ——————————→ 10^{3}

10^{-12} 10^{-9} 10^{-6} 10^{-3} 10^{0}

2nd Best alpha

10^{-9} 10^{-8} -7 -6 -5 -4 10^{-3}

10^{-8} ———— -7 ——

—— —— 10 —— —— 10^{-6}

Train

Brain → 80% NCERT

1b 16 16 1b 16

Train

Mock Test

20% PreBoard → %

R^2

adj

Board

↑

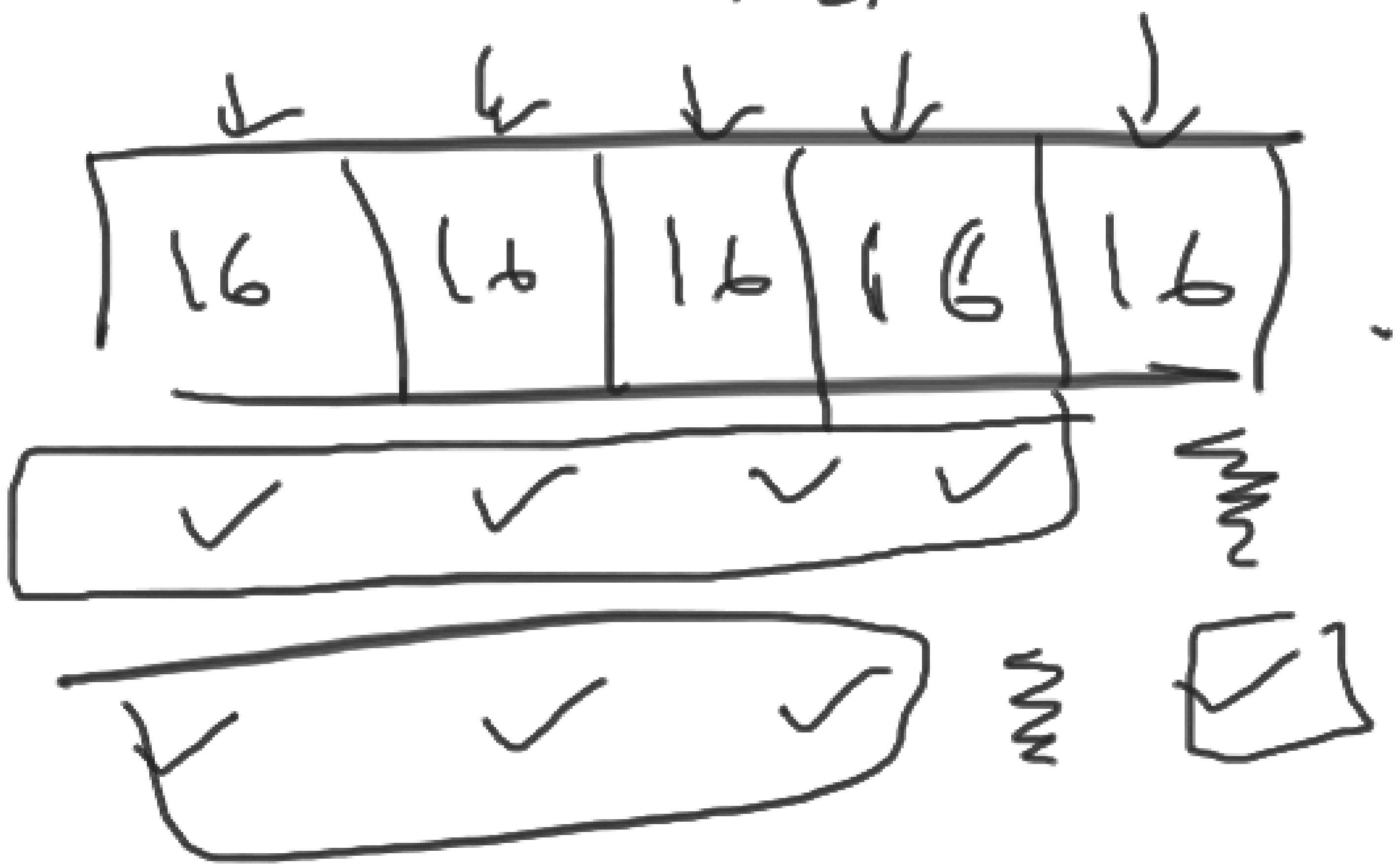
LIVE DATA

Cross-Valid

5 iters

iter 1

2



overfit → avg of ans

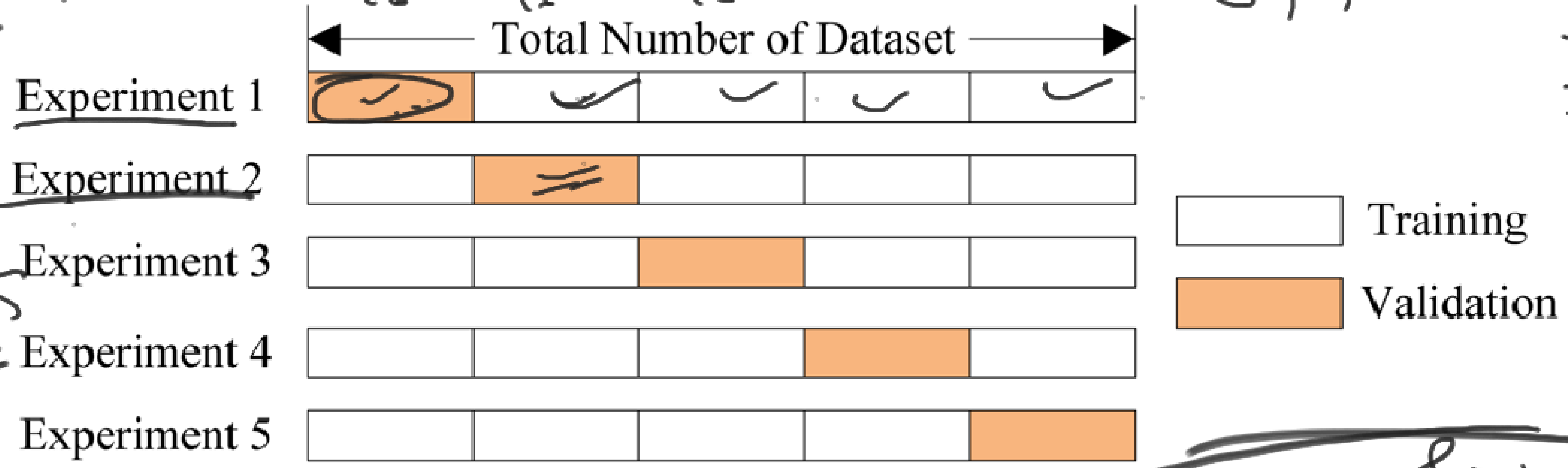
seed

80%

16 16 16 16 16

64%

80
-16



CV=5

overfit

Exp → diff → train → avg -acc

Cross-validation

Gridsearch CV

Reg → Data procure

↳ clean junk → duplicate

→ Nan/missing

→ ~~outliers~~

→ ~~skew~~

EDA → Analyzing → corr

→ unwanted

→ hypothesis

→ encoding

→ outlier

→ skew

featureeng → business
→ VIF → reduce
→ featuretools → adv
golden

Board

test-train → 80% — 80% PB
evaluation

scaling
norma

class → model
→ gridsearchCV (— , param dict)
. fit (X-train, Y-train)
. score → R^2

(MAPE
RMSLE

Best fit → Best param

→neg mean sq

(Med) →

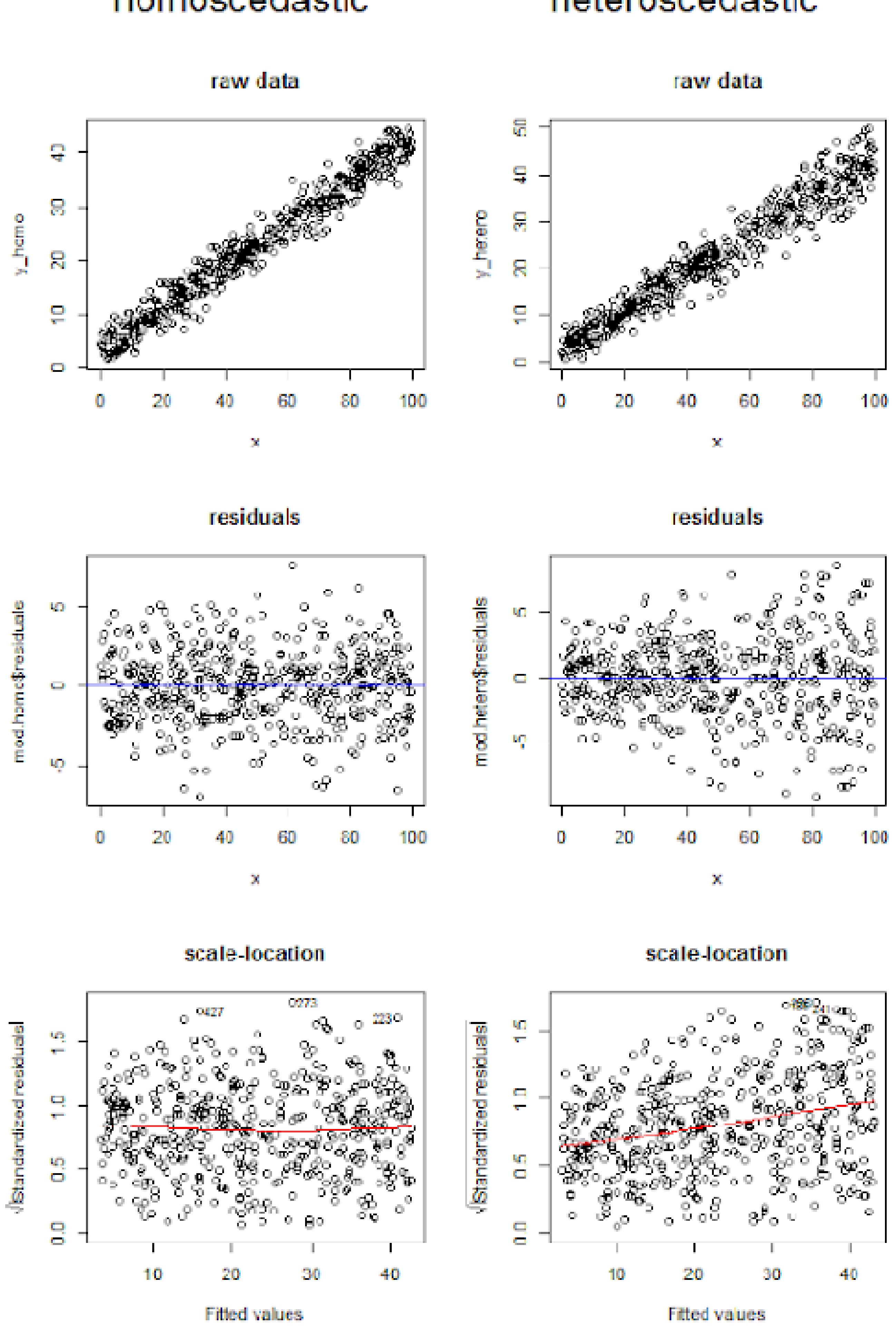
Best EM → (R^2)

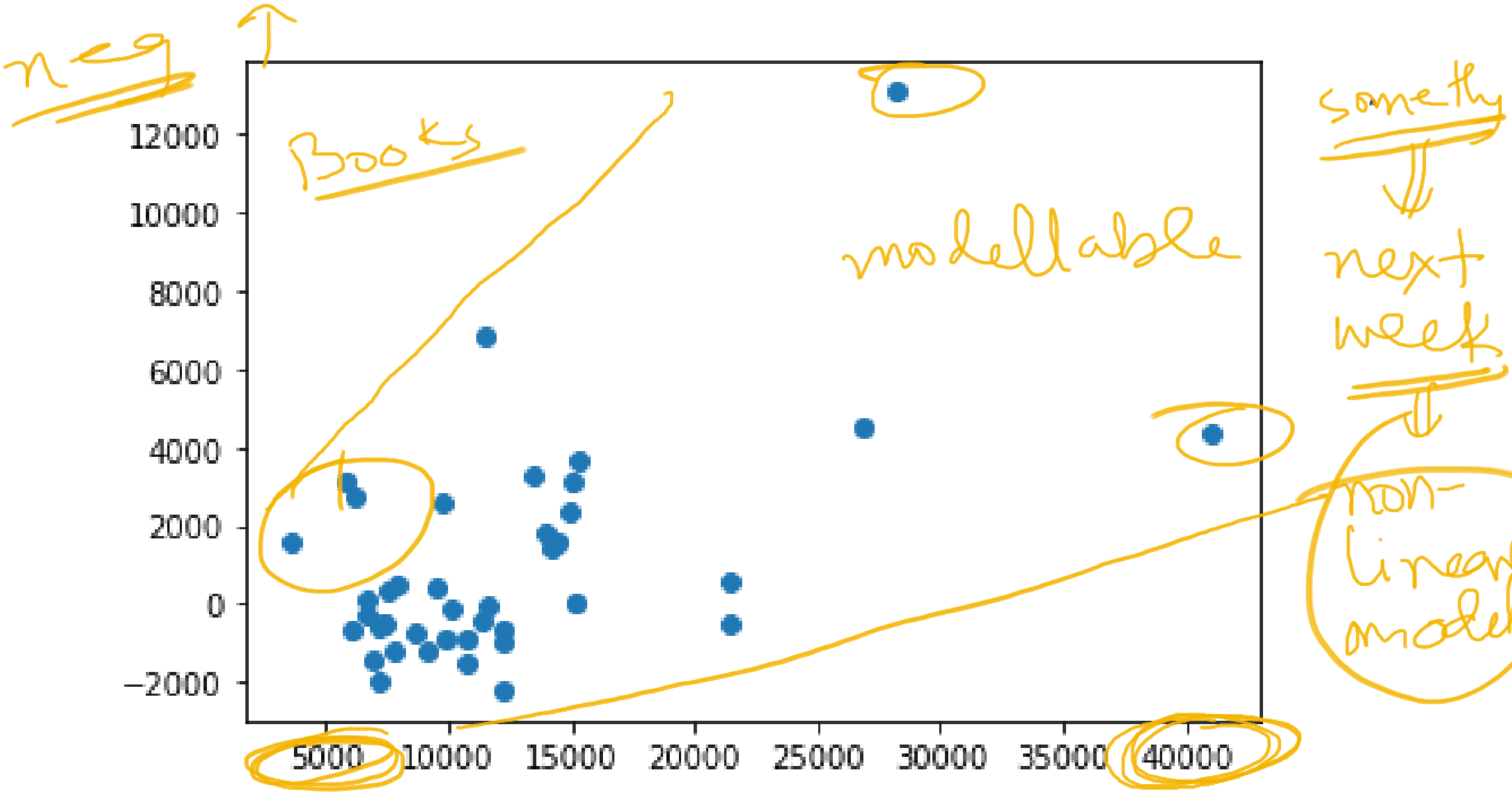
"R^2" → goodness of fit

adj R^2 →

MSE : 8476786.746703872
RMSE : 2911.4921855817975
R2 : 0.8905039330011489
Adjusted R2 : 0.7262598325028724

MSE : 8854606.545200942
RMSE : 2975.669092019632
R2 : 0.8856235717031826
Adjusted R2 : 0.7140589292579566


homoscedastic
heteroscedastic
raw data
raw data
y_homo
y_hetero
x
x
residuals
residuals
mod.homo$residuals
mod.hetero$residuals
x
x
scale-location
scale-location
√|Standardized residuals|
√|Standardized residuals|
Fitted values
Fitted values


neg
Books
modellable
next
week
non-
linear
12000
10000
8000
6000
4000
2000
0
−2000
5000
10000
15000
20000
25000
30000
35000
40000

sns.residplot(x=y_pred, y=y_test, lowess=True, color="g")