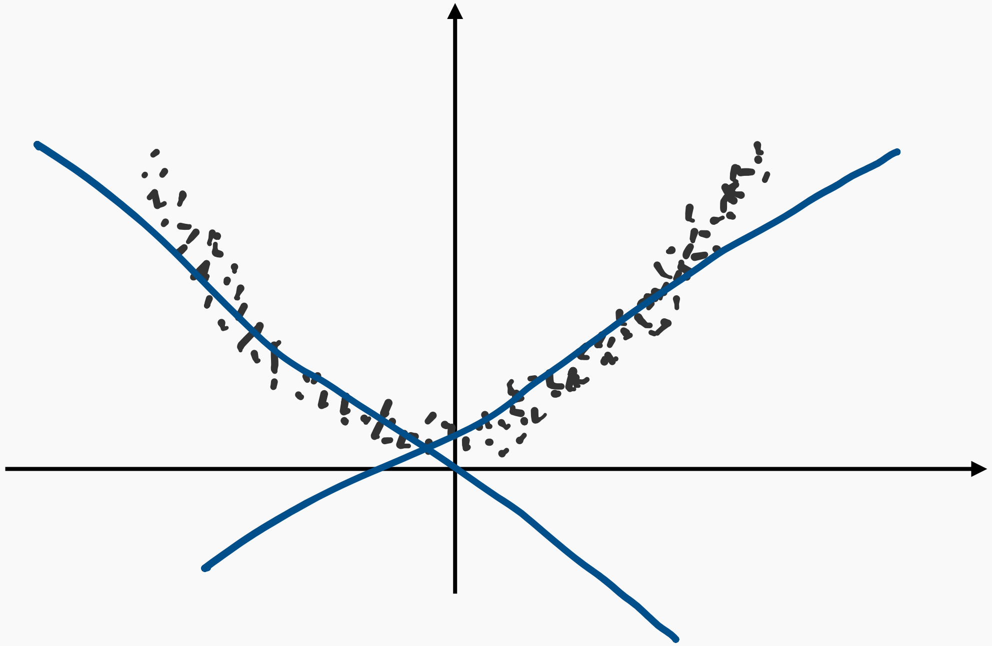
kokchun giang

the art of finding a model not too complex to **overfit** and not too simple to **underfit**

how to fit this data?



- non linear relation
- a line is too simple to capture datas variance

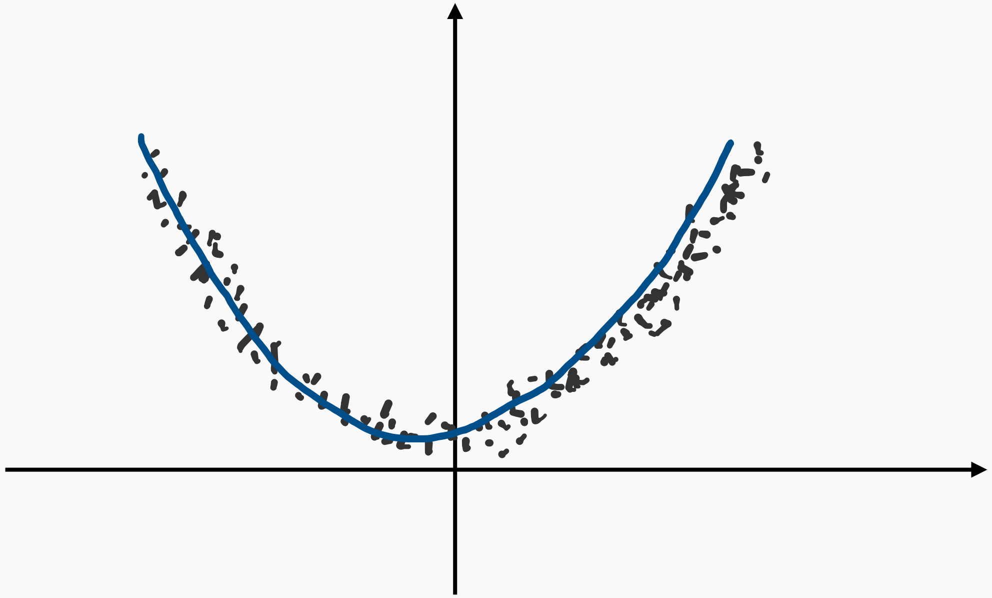
→ underfitting

→ model low attention to training data

↑ bias

↓ variance

feature engineering for polynomial regression



↓ bias

↑ variance

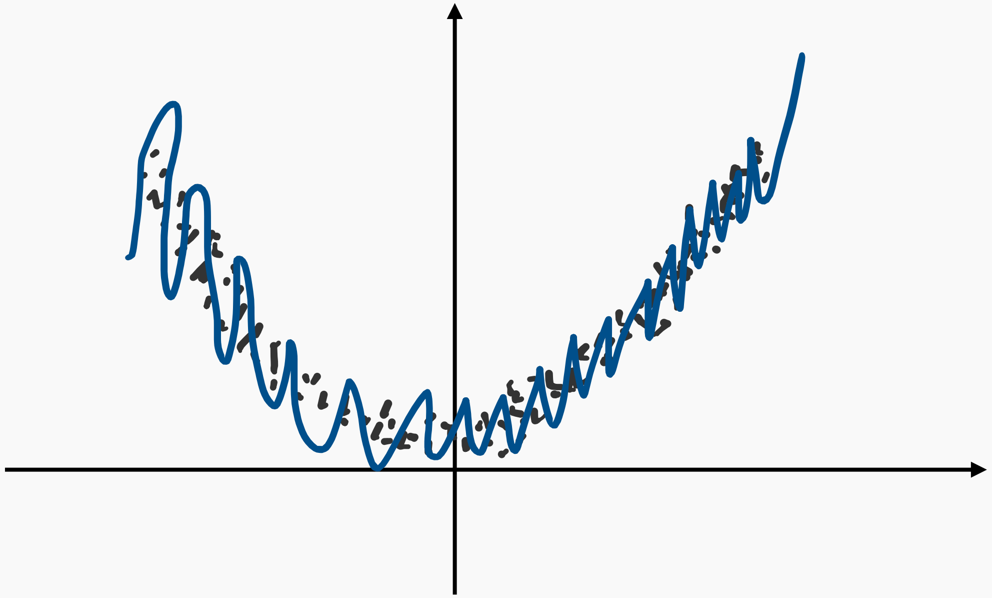
- transform feature space into 2nd degree

→ polynomial regression

ex $\begin{pmatrix} | \\ x_1 \\ | \end{pmatrix} \rightarrow \begin{pmatrix} | & | \\ x_1 & x_1^2 \\ | & | \end{pmatrix}$

2nd order polynomial $a x_1 + b x_1^2$

what if we make a very **complex model**?



↓ bias

↑ variance

- very high degree
- ↓ error in training data
- fits to noise

→ overfitting

- can't generalize well

need to find a balance – **bias-variance tradeoff**

bias - difference between avg prediction and correct value

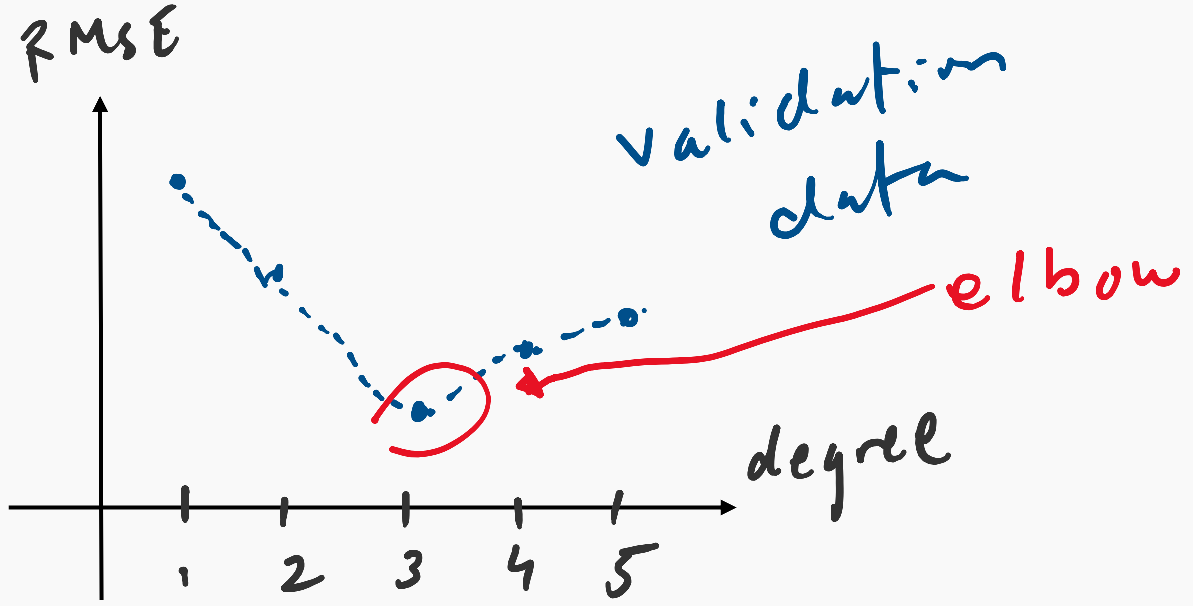
variance - how much model changes if training on multiple datasets

↑ bias
↓ variance → oversimplify model → underfitting

↓ bias
↑ variance → fits to noise → overfitting

elbow plot

$$\text{error} = \underbrace{\text{bias} + \text{variance}}_{\text{our focus}} + \text{irreducible error}$$



here we pick degree 3 for our polynomial regression