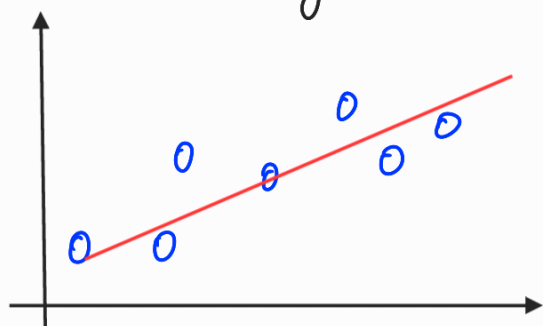


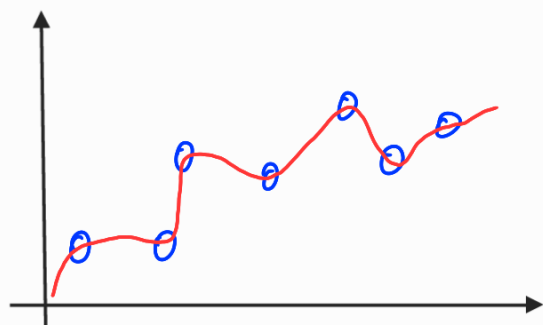
Bias: inability to capture true relationship (poorly).



Low bias: adapts better to new data, gets the relationship.

High bias: does not get relationship.

Variance: inability to capture true relationship (too much).



Seems to fit perfectly for training.

High variance: changes a lot with new data.

Ideally → Low bias: gets the relationship
→ Low variability: consistent predictions

Regularization
Boosting
Bagging

REGULARIZATION Used to get simpler hypothesis.

Cost function + $\lambda \sum_{j=1}^n \theta_j^2$

→ regularization parameter:

Too large: $\theta \approx 0$; underfits

Too small: Does not change