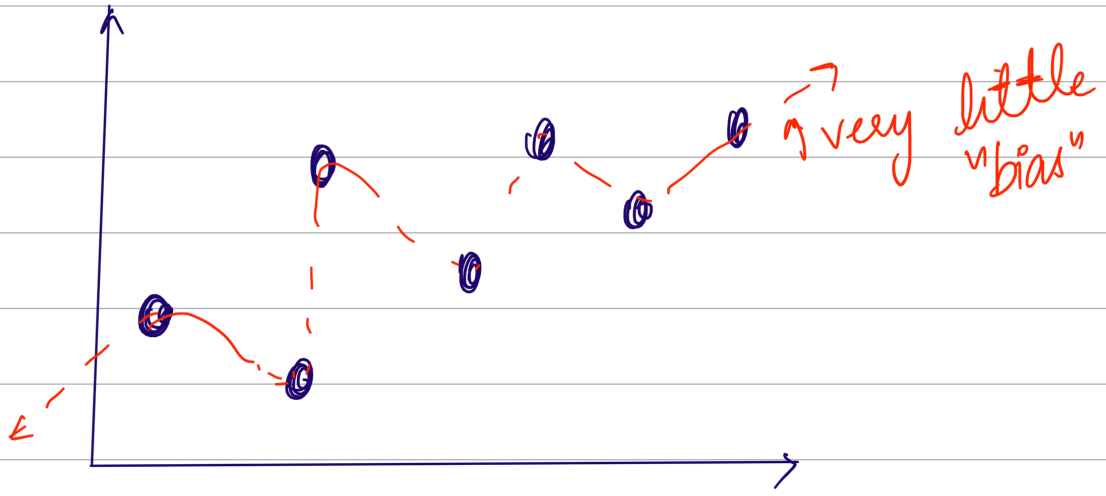
ML algorithm - Linear Regression [Least Squares]

When we use linear regression, it fits a straight line to the training set

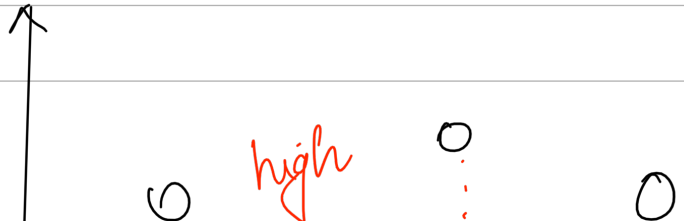
↳ The straight line does not have the flexibility to accurately replicate the arc in the "true" relationship

The inability for a ML method (like linear regression) to capture the true relationship is called bias

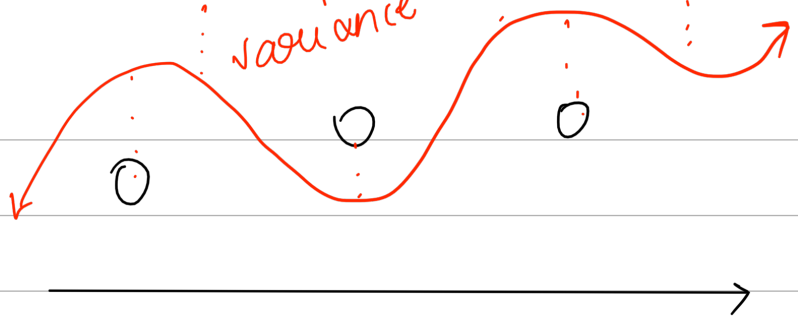


"Sq" distances :- +ve distances do not cancel out -ve distances

Variance - The difference in fits between data sets



Squiggly line is overfit



Ideal alg: low bias → accurately model the true relationship
low variability
↳ produce consistent predictions across different datasets

Find the balance between simple, complex model can be done by methods such as regularization, boosting and bagging