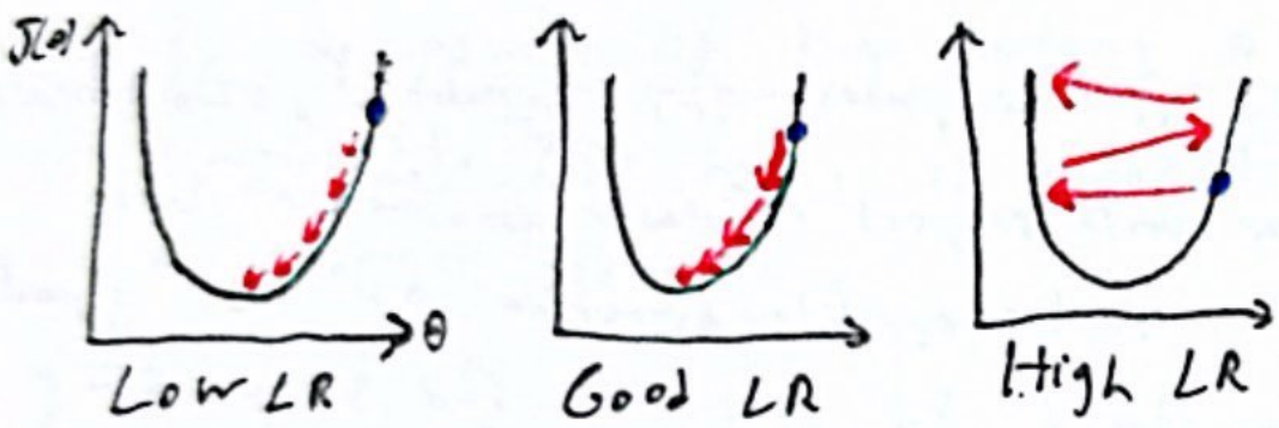
Learning Rate

- is a important Hyper parameter that determines how much a model adjusts its parameters in response to errors during training.

* a model with a hight LR makes Large adjustments to its paramets after each Batch of Data, potentially learning Quickly, but risking overshooting optimal values

* a Low LR means the model learns slower as it makes smaller cautious adjustments, this can be more stable but might take longer to converge or get stuck in suboptimal Solutions

* finding the right LR is important.



Common LR values

Model type	start value
Simple ML models	0.01 or 0.001
NN (basic)	0.001
CNN RNN, gate	0.001 or 0.0001
Transformer/DeepNets	0.0001
When using ADAM	0.001

* 0.1 → too high, Do trail and error

Evaluation

- is the process of Measuring how well a ML model preforms on data is has not seen in training (test set) using metrics appropriate to task. But Evaluation typically involves Validation to tune model in training and testing using test set

Ex1) ////// Preformance metrics //////

classification	Regression	clustering
•Accuracy. % of correct predictions	•MSE (Mean squared error). penalize large errors	•silhouette score - measures how well data points fit into its own clusters vs others •Range (-1, 1) +1: well clustered 0: on boundary -1: misclassified Higher = Better
•percision. TP/TP+FP how many predicted positives correct	•RMSE like MSE but in original units	
•Recall (sensitivity) TP/TP+FN How many actual positives where found	•MAE (Mean absolute error) avg abs error	
•F1 score. Harmonic mean of percition and recall	•R^2 score. % of variance explained by the model	•Davies-Bouldin id.
•ROC - AUC trade off Between TPR and FPR	classification	•Calinski-Harabasz score
	Confusion matrix = table of TP, TN, FP, FN	•ARI - labels are needed

Ex2) Cross validation (see p13)

Ex3) Hold out method